A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations

J.Fontbona* B.Jourdain†

July 19, 2011

Abstract

We introduce and develop a pathwise description of the dissipation of general convex entropies for continuous time Markov processes, based on simple backward martingales and convergence theorems with respect to the tail sigma field. The entropy is in this setting the expected value of a backward submartingale. In the case of (non necessarily reversible) Markov diffusion processes, we use Girsanov theory to explicit its Doob-Meyer decomposition, thereby providing a stochastic analogue of the well known entropy dissipation formula, valid for general convex entropies (including total variation). Under additional regularity assumptions, and using Itô calculus and ideas of Arnold, Carlen and Ju [2], we obtain a new Bakry Emery criterion which ensures exponential convergence of the entropy to 0. This criterion is non-intrisic since it depends on the square root of the diffusion matrix, and cannot be written only in terms of the diffusion matrix itself. Last, we provide an example where the classic Bakry Emery criterion fails, but our non-intrisic criterion ensuring exponential convergence to equilibrium applies without modifying the law of the diffusion process.

Introduction

We are interested in the long-time behaviour of solutions to the stochastic differential equation

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt \tag{0.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\otimes d'}$ and $W = (W_t, t \geq 0)$ is a standard Brownian motion in $\mathbb{R}^{d'}$.

We consider a convex function $U : [0,\infty) \to \mathbb{R}$ bounded from below and define the $U-$entropy of a probability measure p in $\mathbb{R}^d$ with respect to a probability measure q by

$$H_U(p|q) = \begin{cases} \int_{\mathbb{R}^d} U\left(\frac{dp}{dq}(x)\right) dq(x) & \text{if } p \ll q \\ +\infty \text{ otherwise.} \end{cases}$$

The particular cases $U(x) = x\ln(x)\mathbf{1}_{x\geq 0} + (+\infty)\mathbf{1}_{x<0}$ and $U(x) = (x-1)^2$ respectively correspond to the usual entropy and the χ^2-distance. For $U(x) = |x-1|$, $H_U(p|q)$ coincides with the total variation distance but only when $p \ll q$.

*DIM-CMM, UMI(2807) UCHILE-CNRS, Universidad de Chile, Casilla 170-3, Correo 3, Santiago-Chile, e-mail:fontbona@dim.uchile.cl. Supported by Fondecyt Grant 1110923 and BASAL-Conicyt.

†Université Paris-Est, CERMICS, 6-8 av Blaise Pascal, Cité Descartes, Champs sur Marne, 77455 Marne-la-Vallée Cedex 2, France - e-mail : jourdain@cermics.enpc.fr. Supported by ANR-09-BLAN-0216-01 MEGAS.

The celebrated Bakry Emery curvature dimension criterion which involves the generator, the carré du champs and the iterated carré du champs of a continuous-time Markov process is a sufficient condition for the reversible measure of this Markov process to satisfy a Poincaré inequality and a logarithmic Sobolev inequality. From these inequalities, one can respectively deduce exponential convergence to 0 as $t \to \infty$ of the chi-square distance or the relative entropy between the marginal at time t of the process and its reversible measure. This criterion has been generalized to entropy functions U more general than $U(r) = (r-1)^2$ and $U(r) = r\ln(r)$ (see for instance [1]).

In general, even when the stochastic differential equation (0.1) admits an invariant probability measure, this measure is not reversible. It is well known both from a probabilistic point of view [7] and the point of view of partial differential equations [2] that the presence of a contribution antisymmetric with respect to the invariant measure in the drift may accelerate convergence to this invariant measure as $t \to \infty$.

The primal goal of this work was to recover the results of [2] and [1] about the long-time behaviour of U-entropy of the law of X_t with respect to the invariant measure by arguments based on Itô's stochastic calculus. To achieve this goal, we introduce and develop in the first section of the paper a pathwise description of the dissipation of general convex entropies for continuous time non-homogeneous Markov processes, based on simple backward martingales and convergence theorems with respect to the tail sigma field. Given two different initial laws, the U-entropy of the marginal at time t of the Markov process under the first initial law with respect to its marginal at time t under the second initial law is the expected value of a backward submartingale. This implies that this U-entropy is non-increasing with t and permits to characterize its limit as $t \to \infty$. To our knowledge, this simple result does not seem to have been used in the study of the trend to equilibrium of Markov processes.

From the second section of the paper on, we only deal with Markov diffusions given by

$$dX_t = b(t, X_t) + \sigma(t, X_t)dW_t \tag{0.2}$$

where $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \otimes d'}$. Under assumptions that guarantee that for both initial laws, the time-reversed processes are still diffusions, we use Girsanov theory to explicit the Doob-Meyer decomposition of the submartingale obtained in the first section. In this way, we obtain a stochastic analogue of the well known entropy dissipation formula, valid for general convex entropies (including total variation). Taking expectations in this formula, we recover the well known fact that the U-entropy dissipation is equal to the U-Fisher information.

It should be noticed that the idea of considering a trajectorial interpretation of entropy to obtain functional inequalities is not new, at least for reversible diffusions (see e.g. the work of Cattiaux [3] whose results nevertheless are of quite different nature). However, even in the reversible case, time reversal of a diffusion starting out of equilibrium modifies the dynamics of the diffusion. The simple martingales introduced in the first section take this fact into account and moreover permit the use of Itô calculus under less regularity than *a priori* needed when working in the forward time direction. Their interest thus goes beyond the treatment of non-reversible situations.

In the third section, we further suppose that the stochastic differential equation is time-homogeneous and that it admits an invariant probability distribution, that is chosen as one of the two initial laws. Under additional regularity assumptions, and using Itô calculus and ideas of Arnold, Carlen and Ju [2], we obtain a new Bakry Emery criterion which ensures exponential convergence of the U-Fischer information to 0 and therefore exponential convergence of the U-entropy to 0. In addition, under this criterion, the invariant measure satisfies a U-convex Sobolev inequality. This criterion is non-intrisic : it depends on the square root σ of the diffusion matrix $a = \sigma\sigma^*$

and cannot be written only in terms of the diffusion matrix itself whereas, under mild regularity assumptions on b and a, the law of $(X_t)_{t\geq 0}$ solving (0.1) is characterized by the associated martingale problem only written in terms of a and b. Last, we provide an example where the classic Bakry Emery criterion fails, but our non-intrisic criterion ensures exponential convergence to equilibrium without modifying the law of the diffusion process. As future work, we plan to investigate how to choose the square root σ of the diffusion matrix in order to maximize the rate of exponential convergence to equilibrium given by our non-intrisic Bakry Emery criterion.

Acknowledgements : We thank Tony Lelièvre (CERMICS) for pointing out to us the paper of Arnold, Carlen and Ju [2] at an early stage of this research.

1 Entropy decrease for general continuous-time Markov processes

Throughout this work, we make the following assumption on U:

$H0)$ $U : [0, \infty) \to \mathbb{R}$ is a convex function such that $\inf U > -\infty$.

Notice that U is then continuous on $(0, +\infty)$ and such that $U(0) \geq \lim_{x\to 0^+} U(x)$.

In this section it is assumed that $(X_t : t \geq 0)$ is a continuous-time Markov process with values in $(E, \mathcal{E})$.

Let P_0, Q_0 be probability measures on E. We will use throughout the following notation:

- $(X_t^{P_0}, t \geq 0)$ and $(X_t^{Q_0}, t \geq 0)$ denote realizations of the process (X_t) with $X_0^{P_0}$ and $X_0^{Q_0}$ respectively distributed according to P_0 and Q_0.
- For each $t > 0$, P_t and Q_t then stand for the laws of $X_t^{P_0}$ and $X_t^{Q_0}$, respectively.

Proposition 1.1 *The function $t \in \mathbb{R}_+ \mapsto H_U(P_t|Q_t) \in \mathbb{R} \cup \{+\infty\}$ is non-increasing.*

Moreover, if for some $t \geq 0$, $P_t \ll Q_t$, then the law of $(X_r^{P_0})_{r\geq t}$ is absolutely continuous with respect to the one of $(X_r^{Q_0})_{r\geq t}$ with density $\frac{dP_t}{dQ_t}(X_t^{Q_0})$, for all $s \geq t$ it holds that $P_s \ll Q_s$, and $\left(\frac{dP_s}{dQ_s}(X_s^{Q_0})\right)_{s\geq t}$ is a backward martingale with respect to the filtration $\mathcal{F}_s = \sigma(X_r^{Q_0}, r \geq s)$.

Last, if $H_U(P_t|Q_t) < +\infty$ for some $t \geq 0$, then $\left(U(\frac{dP_s}{dQ_s}(X_s^{Q_0}))\right)_{s\geq t}$ is a backward submartingale with respect to $\mathcal{F}_s$.

Corollary 1.2 *If $H_U(P_t|Q_t) < +\infty$ for some $t \geq 0$, then*

$$\lim_{s\to\infty} H_U(P_s|Q_s) = \mathbb{E}\left(U\left(\lim_{s\to\infty} \frac{dP_s}{dQ_s}(X_s^{Q_0})\right)\right) < \infty.$$

In particular, if $U(1) = 0$ and the tail σ-field $\cap_{s\geq 0}\mathcal{F}_s$ is trivial a.s. then $\lim_{s\to\infty} H_U(P_s|Q_s) = 0$.

Proof of Proposition 1.1. Let $s \geq t \geq 0$. If P_t is not absolutely continuous with respect to

Q_t, then $+\infty = H_U(P_t|Q_t) \geq H_U(P_s|Q_s)$. Otherwise, for $f : E^{\mathbb{R}_+} \to \mathbb{R}$ measurable with respect to the product sigma-field and $\mathbb{E}_{t,x}$ the conditional expectation given $X_t = x$, one has

$$\begin{aligned}\mathbb{E}(f(X_r^{P_0}, r \geq t)) &= \int_{\mathbb{R}^d} \mathbb{E}_{t,x}(f(X_r, r \geq 0))P_t(dx) = \int_{\mathbb{R}^d} \mathbb{E}_{t,x}\left(f(X_r, r \geq 0)\frac{dP_t}{dQ_t}(X_0)\right) Q_t(dx)\\ &= \mathbb{E}(f(X_r^{Q_0}, r \geq t)\frac{dP_t}{dQ_t}(X_t^{Q_0})). \end{aligned} \tag{1.1}$$

Hence the law of $(X_r^{P_0})_{r\geq t}$ is absolutely continuous with respect to the one of $(X_r^{Q_0})_{r\geq t}$ with density $\frac{dP_t}{dQ_t}(X_t^{Q_0})$ and $\forall r \geq t$, $P_r \ll Q_r$. Now, for $s \geq t$,

$$\mathbb{E}\left(f(X_r^{Q_0}, r \geq s)\frac{dP_t}{dQ_t}(X_t^{Q_0})\right) = \mathbb{E}\left(f(X_r^{P_0}, r \geq s)\right) = \mathbb{E}\left(f(X_r^{Q_0}, r \geq s)\frac{dP_s}{dQ_s}(X_s^{Q_0})\right)$$

where we used (1.1) with t replaced by s. This ensures that $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})\middle|\mathcal{F}_s\right) = \frac{dP_s}{dQ_s}(X_s^{Q_0})$. By Jensen's inequality, since U is a convex function bounded from below,

$$\mathbb{E}\left(U\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})\right)\middle|\mathcal{F}_s\right) \geq U\left(\frac{dP_s}{dQ_s}(X_s^{Q_0}))\right). \tag{1.2}$$

Taking expectations one concludes that $H_U(P_t|Q_t) \geq H_U(P_s|Q_s)$. ∎

Proof of Corollary 1.2. If $H_U(P_t|Q_t) < +\infty$ then $P_t \ll Q_t$ and the $\mathcal{F}_s$ backward martingale $(\frac{dP_s}{dQ_s}(X_s^{Q_0}))_{s\geq t}$ converges a.s. to $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})\middle|\cap_{s\geq 0}\mathcal{F}_s\right)$ when $s \to \infty$. By the backward martingale property, for $r \geq t$,

$$\mathbb{E}\left(\frac{dP_r}{dQ_r}(X_r^{Q_0})1_{\{\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})|\cap_{s\geq 0}\mathcal{F}_s\right)=0\}}\right) = \mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})1_{\{\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})|\cap_{s\geq 0}\mathcal{F}_s\right)=0\}}\right) = 0.$$

Hence $\frac{dP_r}{dQ_r}(X_r^{Q_0}) = 0$ a.s. on the set $\left\{\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})\middle|\cap_{s\geq 0}\mathcal{F}_s\right) = 0\right\}$. With the continuity of U on $(0, +\infty)$, one deduces that the random variables $U\left(\frac{dP_s}{dQ_s}(X_s^{Q_0})\right)$ converge a.s. to $U\left(\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})\middle|\cap_{s\geq 0}\mathcal{F}_s\right)\right)$ as $s \to +\infty$. Since they are uniformly integrable, one concludes that $H_U(P_s|Q_s) = \mathbb{E}\left(U\left(\frac{dP_s}{dQ_s}(X_s^{Q_0})\right)\right)$ converges as $s \to \infty$ to the asserted limit. When the tail σ-field is trivial a.s., the limit of the backward martingale is equal to $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^{Q_0})\right) = 1$ and $U\left(\frac{dP_s}{dQ_s}(X_s^{Q_0})\right)$ converges a.s. to $U(1)$. ∎

2 Entropy dissipation for diffusion processes

From now on we assume that $(X_t, t \geq 0)$ is a Markov diffusion process solution to equation (0.2).

We introduce a finite time-horizon $T \in (0, +\infty)$ in order to define a forward martingale associated with $\frac{dP_s}{dQ_s}(X_s^{Q_0})$ by time-reversal. Moreover,

- we denote by $\mathbb{P}^T$, $\mathbb{Q}^T$, $\mathbb{P}^{T\to 0}$ and $\mathbb{Q}^{T\to 0}$ the respective laws of $(X_t^{P_0}, t \leq T)$, $(X_t^{Q_0}, t \leq T)$, $(\bar{X}_t^{P_0,T} := X_{T-t}^{P_0}, t \leq T)$ and $(\bar{X}_t^{Q_0,T} := X_{T-t}^{Q_0}, t \leq T)$ on $C([0,T], \mathbb{R}^d)$ and by $\mathbb{E}^T$, $\tilde{\mathbb{E}}^T$, $\mathbb{E}^{T\to 0}$ and $\tilde{\mathbb{E}}^{T\to 0}$ the corresponding expectations.

- We also denote by $(Y_t)_{t\le T}$ the canonical process on $C([0,T],\mathbb{R}^d)$ and by $\mathcal{G}_t = \sigma(Y_s, 0 \le s \le t)$ its filtration.
- For $0 \le t \le T$, let finally $\mathcal{H}_t^T := \sigma(\bar{X}_s^{Q_0,T}, 0 \le s \le t) = \sigma(X_s^{Q_0}, T-t \le s \le T)$.

Lemma 2.1 *If $P_0 \ll Q_0$, then $\mathbb{P}^{T\to 0} \ll \mathbb{Q}^{T\to 0}$, $\frac{d\mathbb{P}^{T\to 0}}{d\mathbb{Q}^{T\to 0}} = \frac{dP_0}{dQ_0}(Y_T)$ and $\frac{dP_{T-t}}{dQ_{T-t}}(\bar{X}_t^{Q_0,T})$, $0 \le t \le T$ is a (uniformly integrable) backward martingale with respect to $\mathcal{H}_t^T$. Last, on the canonical space*

$$D_t^T \overset{\text{def}}{=} \frac{d\mathbb{P}^{T\to 0}}{d\mathbb{Q}^{T\to 0}} |_{\mathcal{G}_t} = \frac{dP_{T-t}}{dQ_{T-t}}(Y_t), \quad 0 \le t \le T,$$

is a $\mathbb{Q}^{T\to 0} - \mathcal{G}_t$ martingale with a right continuous version also denoted D_t^T.

Proof of Lemma 2.1. Since $\mathbb{P}^{T\to 0}$ and $\mathbb{Q}^{T\to 0}$ are the respective images of $\mathbb{P}^T$ and $\mathbb{Q}^T$ by time-reversal : $(Y_s, 0 \le s \le T) \to (Y_{T-s}, 0 \le s \le T)$ one deduces that $\mathbb{P}^{T\to 0} \ll \mathbb{Q}^{T\to 0}$ with $\frac{d\mathbb{P}^{T\to 0}}{d\mathbb{Q}^{T\to 0}} = \frac{dP_0}{dQ_0}(Y_T)$. Reasoning as in the proof of Proposition 1.1, one checks that $\left(\frac{dP_t}{dQ_t}(X_t^{Q_0}) = \frac{dP_t}{dQ_t}(\bar{X}_{T-t}^{Q_0,T})\right)_{0\le t\le T}$ is a backward martingale with respect to the filtration $\mathcal{H}_{T-t}^T$. Therefore by time-reversal, $\left(\frac{dP_{T-t}}{dQ_{T-t}}(\bar{X}_t^{Q_0,T})\right)_{0\le t\le T}$ is a martingale with respect to the filtration $\mathcal{H}_t^T$. On the canonical space, one deduces that

$$\tilde{\mathbb{E}}^{T\to 0}\left(\frac{d\mathbb{P}^{T\to 0}}{d\mathbb{Q}^{T\to 0}}\middle|\mathcal{G}_t\right) = \tilde{\mathbb{E}}^{T\to 0}\left(\frac{dP_0}{dQ_0}(Y_T)\middle|\mathcal{G}_t\right) = \frac{dP_{T-t}}{dQ_{T-t}}(Y_t).$$

∎

Remark 2.2 *a) Similar arguments as in* (1.2) *shows that for each $s \in [0,T]$, $H_U(P_s|Q_s) < +\infty$ if and only if $U\left(\frac{dP_{T-t}}{dQ_{T-t}}(Y_t)\right), 0 \le t \le T-s$, is a uniformly integrable $\mathbb{Q}^{T\to 0}$ submartingale with respect to $\mathcal{G}_t$.*

b) Given two probability measure $\mathbb{P}_1, \mathbb{P}_2 \in C([0,T],\mathbb{R}^d)$, the pathwise $U-$entropy defined by

$$\mathbb{H}_U(\mathbb{P}_1|\mathbb{P}_2) = \begin{cases} \int_{C([0,T],\mathbb{R}^d)} U\left(\frac{d\mathbb{P}_1}{d\mathbb{P}_2}(w)\right) d\mathbb{P}_2(w) & \text{if } \mathbb{P}_1 \ll \mathbb{P}_2 \\ +\infty & \text{otherwise,} \end{cases}$$

satisfies $H_U(P_0|Q_0) = \mathbb{H}_U(\mathbb{P}^T|\mathbb{Q}^T) = \mathbb{H}_U(\mathbb{P}^{T\to 0}|\mathbb{Q}^{T\to 0})$, thanks to Lemma 2.1.

Since for $0 \le t \le T$, $H_U(P_t|Q_t) = \mathbb{E}\left(U\left(\frac{dP_t}{dQ_t}(\bar{X}_{T-t}^{Q_0,T})\right)\right)$, to precise how this quantity decreases in time we will be interested in the $\mathcal{H}_t^T$-martingale $\left(\frac{dP_{T-t}}{dQ_{T-t}}(\bar{X}_t^{Q_0,T}), 0 \le t \le T\right)$ and therefore in the time-reversal $(\bar{X}_t^{Q_0,T}, 0 \le t \le T)$ of the diffusion process $(X_t^{Q_0}, 0 \le t \le T)$.

We assume from now on that the Markov process $\bar{X}_t^{Q_0,T}$ is again a diffusion process. Conditions ensuring this fact have been studied among other authors by Föllmer [4], Hausmann and Pardoux [6], Pardoux [12] and Millet *et. al* [11], who in particular provide the semimartingale decomposition of $\bar{X}_t^{Q_0,T}$ in its filtration. We shall base ourselves on the general results in [11], which we recall in Theorem 2.3 below in a slightly more restrictive situation, and which rely on the following conditions:

$H1)$ For each $T > 0$, $\sup_{t\in[0,T]}(|b(t,0)| + |\sigma(t,0)|) < +\infty$ and there exist $K_T > 0$ such that

$$\forall t \in [0,T],\ \forall x,y \in \mathbb{R}^d,\ |b(t,x) - b(t,y)| + \sum_{i=1}^{d'} |\sigma_{\bullet i}(t,x) - \sigma_{\bullet i}(t,y)| \le K_T|x-y|,$$

where $\sigma_{\bullet i}$ denotes the i-th column of the matrix σ.

$H2)_{Q_0}$ For each $t > 0$, the law $Q_t(dx)$ of $X_t^{Q_0}$ has a density $q_t(x)$ with respect to Lebesgue measure.

$H3)_{Q_0}$ Denoting $a_{ij} = (\sigma\sigma^*)_{ij}$, for each $i = 1,\dots,d$ the distributional derivative $\partial_j(a_{ij}(t,x)q_t(x))$ (with summation over repeated indexes) is a locally integrable function on $[0,T] \times \mathbb{R}^d$:

$$\int_0^T \int_D |\partial_j(a_{ij}(t,x)q_t(x))|dxdt < \infty \text{ for any bounded open set } D \subset \mathbb{R}^d.$$

We set for $(t,x) \in [0,T] \times \mathbb{R}^d$

- $\bar{a}_{ij}(t,x) := a_{ij}(T-t,x), i,j = 1,\dots,d,$
- $\bar{b}^i_{Q_0}(t,x) = -b^i(T-t,x) + \frac{\partial_j(a_{ij}(T-t,x)q_{T-t}(x))}{q_{T-t}(x)}$ (with the convention that the term involving $q_{T-t}(x)^{-1}$ is 0 if $q_{T-t}(x)$ is 0)

and notice that $\bar{b}_{Q_0}(t,x)$ is defined $dt \otimes dx$ a.e. on $[0,T] \times \mathbb{R}^d$ under assumption $H3)_{Q_0}$.

Theorem 2.3 *Assume that $H1)$ and $H2)_{Q_0}$ hold.*

a) Suppose moreover that $H3)_{Q_0}$ holds. Then, $\mathbb{Q}^{T\to 0}$ is a solution to the martingale problem:

$$(MP)_{Q_0}: \quad M_t^f := f(Y_t) - f(Y_0) - \int_0^t \frac{1}{2}\bar{a}_{ij}(s,Y_s)\partial_{ij}f(Y_s) + \bar{b}^i_{Q_0}(s,Y_s)\partial_i f(Y_s)ds,\ t \in [0,T]$$

is a continuous martingale with respect to the filtration $(\mathcal{G}_t)$ for all $f \in C_0^\infty(\mathbb{R}^d)$.

b) Let $\tilde{b} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\tilde{\sigma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\otimes d'}$ be measurable functions such that $\int_0^T \int_D |\tilde{a}_{ij}(t,x)| + |\tilde{b}^i(t,x)|q_{T-t}(x)dxdt < \infty$ for any bounded open set $D \subset \mathbb{R}^d$. Assume moreover that $\mathbb{Q}^{T\to 0}$ is a solution to the martingale problem w/r to $(\mathcal{G}_t)$ for the generator $\mathcal{L}_t f(x) = \frac{1}{2}\tilde{a}_{ij}(t,x)\partial_{ij}f(x) + \tilde{b}^i(t,x)\partial_i f(x)$. Then, $\tilde{b} = \bar{b}$, $\tilde{a} = \bar{a}$ and $H3)_{Q_0}$ holds.

Notice when f is C^∞ on $\mathbb{R}^d$ and vanishes outside the ball $B(0,A)$ that

$$\begin{aligned}\tilde{\mathbb{E}}^{T\to 0}&\left(\int_0^T |\bar{b}^i_{Q_0}(s,Y_s)||\partial_i f(Y_s)|ds\right) \\ &\le \sup_{B(0,A)} |\nabla f| \left(T \sup_{[0,T]\times B(0,A)} |b(s,x)| + \int_{[0,T]\times B(0,A)} \left|\sum_{i=1}^{d'} \partial_j(a_{ij}(s,x)q_s(x))\right| dsdx\right)\end{aligned} \tag{2.1}$$

where the right-hand-side is finite under $H1)$ and $H3)_{Q_0}$.

Proof . According to Theorem 2.3 [11], under $H1)$, $H2)_{Q_0}$ and $H3)_{Q_0}$ $(M_t^f)_{t\in[0,T)}$ is a continuous $\mathcal{G}_t$-martingale under $\mathbb{Q}^{T\to 0}$. Since by (2.1) and $H1)$, $t \mapsto M_T^f$ is continuous on $[0,T]$ and $\tilde{\mathbb{E}}^{T\to 0}(|M_T^f|) < +\infty$, one deduces that $(M_t^f)_{t\in[0,T]}$ is a continuous $\mathcal{G}_t$-martingale under $\mathbb{Q}^{T\to 0}$. Part b) follows from Theorem 2.2 in [11]. ∎

Remark 2.4 *i) Under $H1$), condition $H3)_{Q_0}$ is implied by the following one*

$C)_{Q_0}$: Q_0 has a density q_0 w.r.t. the Lebesgue measure s.t. $\exists k>0$, $\int_{\mathbb{R}^d} \frac{q_0^2(x)dx}{1+|x|^k} < +\infty$ and either

$$\forall T>0,\ \exists \varepsilon>0,\ \forall (t,x)\in[0,T]\times\mathbb{R}^d,\ a(t,x){=}\sigma\sigma^*(t,x)\geq \varepsilon I_d,$$

or the second order distribution derivatives $\frac{\partial^2 a^{ij}}{\partial x_i \partial x_j}(t,x)$ are bounded on $[0,T]\times\mathbb{R}^d$ for each $T>0$. Indeed, by Theorem 3.1 in [6], $C)_{Q_0}$ ensures condition (A)(ii) in p. 1189 therein, which implies $H3)_{Q_0}$ when $H1$) holds.

ii) By Theorem 3.3 in [11] and the proof of Theorem 2.2 therein (see also p. 220 in [11]), the global Lipschitz assumption $H1$) in the previous result can be replaced by a local one under additional regularity of the coefficients and exponential integrability of their derivatives. The results in this section will be stated under $H1$), but they also hold in that more general setting of [11].

The next result will give almost everywhere sense to the functions required to describe the processes $(D_t^T)_{t\in[0,T]}$, without additional assumptions.

Lemma 2.5 *Assume that $H1$) , $H2)_{P_0}$ and $H3)_{P_0}$ hold.*

i) For each $i=1\dots,d$ and a.e. $t\in(0,T]$, the distribution $a_{ij}(t,\cdot)\partial_j p_t := \partial_j(a_{ij}(t,\cdot)p_t) - p_t\partial_j a_{ij}(t,\cdot)$ is a function in $L^1_{loc}(dx)$ and, as a Radon measure in $[0,T]\times\mathbb{R}^d$, one has $a_{ij}(t,x)\partial_j p_t(x)dx\ dt \ll p_t(x)dx\ dt$. A version of the Radon-Nikodyn density (measurable in (t,x)) is given by $[a_{ij}(t,\cdot)\partial_j p_t](x)/p_t(x)$. Moreover, there exists a measurable function $(t,x)\mapsto K^p(t,x)\in\mathbb{R}^d$ such that for each $i=1\dots,d$

$$[a_{ij}(t,\cdot)\partial_j p_t](x)/p_t(x) = a_{i\bullet}(t,x)^* K^p(t,x),\ \ p_t(x)dx\ dt\ a.e.$$

ii) If moreover $H2)_{Q_0}$, $H3)_{Q_0}$ and $P_0 \ll Q_0$ hold, one has $a_{ij}(t,x)\partial_j p_t(x)dx\ dt \ll q_t(x)dx\ dt$ and $[a_{ij}(t,\cdot)\partial_j p_t](x)/q_t(x)$ is a (measurable in (t,x)) version of the Radon-Nikodyn derivative. Furthermore, it holds $p_{T-t}(x)dx\ dt$ (but not necessarily $q_{T-t}(x)dx\ dt$) a.e. that

$$\begin{aligned}\bar{b}^i_{P_0}(t,x) - \bar{b}^i_{Q_0}(t,x) =& [\bar{a}_{ij}(t,\cdot)\partial_j p_{T-t}](x)/p_{T-t}(x) - [\bar{a}_{ij}(t,\cdot)\partial_j q_{T-t}](x)/q_{T-t}(x)\\ =& \bar{a}_{i\bullet}(t,x)^*(K^p(T-t,x) - K^q(T-t,x)),\end{aligned}$$

and $q_{T-t}(x)dx\ dt$ (and thus $p_{T-t}(x)dx\ dt$) a.e. that

$$\frac{p_{T-t}(x)}{q_{T-t}(x)}(\bar{b}^i_{P_0}(t,x) - \bar{b}^i_{Q_0}(t,x)) = \frac{p_{T-t}(x)}{q_{T-t}(x)}\bar{a}_{i\bullet}(t,x)^*(K^p(T-t,x) - K^q(T-t,x)).$$

Proof . The Lipschitz character of a (following from our assumptions) ensures that a has a.e. defined derivatives in L^∞ and thus that the distribution $a_{ij}(t,\cdot)\partial_j p_t$ as defined is a function in $L^1_{loc}(dx)$ under $H3)_{P_0}$. This implies, by Lemma A.2 in [11] (see also Lemma A.2 in [6]), that $a_{ij}(t,x)\partial_j p_t(x)$ vanishes a.e. on $\{x : p_t(x)=0\}$. This fact easily yields the remaining assertions, except for the existence of the functions K^p or K^q, which we establish in what follows.

We will on one hand use the fact, asserted in the proof of Lemma A.2 in [11], that for each $t>0$ and each bounded open set Θ, $a_{ij}(t,x)\partial_j p_t(x)$ is the $\sigma(L^1(\Theta), L^\infty(\Theta))$–weak limit of some subsequence of $a_{ij}(t,x)\partial_j[\rho_n * p_t](x)$, for rapidly decaying regularizing kernels ρ_n. It is indeed shown in

Lemma A.1 in [6] that for a suitable bounded sequence $\alpha_n > 0$, $\alpha_n^{-1}|x|\,|\nabla\rho_n(x)|$ is again a regularizing kernel. The Lipschitz character of a then yields the domination $|a_{ij}(t,x)\partial_j[\rho_n * p_t](x)| \leq |\rho_n * \partial_j(a_{ij}(t,\cdot)p_t)(x)| + C\alpha_n^{-1}\int|x-y|\,|\nabla\rho_n(x-y)|p_t(y)dy$, the right hand side being, by the previous, an $L^1(\Theta)$-converging sequence. Weak compactness is then provided by the Dunford-Pettis criterion, and the limit is identified integrating by parts against smooth test functions compactly supported in Θ. On the other hand, we will use the fact that diagonalizing the symmetric positive semidefinite matrix $(a_{ij}(t,x)) = [u_1(t,x),\dots,u_d(t,x)]\Lambda(t,x)[u_1(t,x),\dots,u_d(t,x)]^*$ provides orthonormal vectors $(u_i(t,x))_{i=1}^d$ and the corresponding eigenvalues and diagonal components $(\lambda_i(t,x))_{i=1}^d$ of $\Lambda(t,x)$, that are measurable as functions of (t,x).

We take Θ as before and $a_{ij}(t,x)\partial_j[\rho_n * p_t](x)$ to be the subsequence described above. Defining the vectorial functions $w^{(n)} := [u_1,\dots,u_d]^*\nabla[\rho_n * p_t]$ and $v_k = sign(u_k^*[a\nabla p])u_k,\ k=1,\dots,d$, we have

$$\int_{\Theta\cap\{\lambda_k=0\}}|v_k^*[a\nabla p_t]| = \lim_{n\to\infty}\int_{\Theta\cap\{\lambda_k=0\}} v_k^*[a\nabla[\rho_n * p_t]] = \lim_{n\to\infty}\int_{\Theta\cap\{\lambda_k=0\}}\lambda_k w_k^{(n)} sign(u_k^*[a\nabla p_t]) = 0,$$

since $a\nabla[\rho_n * p_t] = \sum_{j=1}^d \lambda_j w_j^{(n)} u_j$ by the spectral decomposition of a. Consequently, for each t and a.e. $x\in\mathbb{R}^d$, the vector $[a(t,x)\nabla p_t(x)]$ belongs to the linear space $\left\langle (u_i(t,x))_{i=1,\dots,d;\lambda_i(t,x)\neq 0}\right\rangle$. Denote now by $w = (w_j)_{j=1}^d := (u_j^* a\nabla p_t)_{j=1}^d$ the coordinates of $a\nabla p_t$ w.r.t. the orthogonal basis $(u_j(t,x))_{j=1,\dots,d}$, so that w is a measurable function of (t,x). If we moreover denote by $\overline{\Lambda}$ the diagonal matrix with diagonal $\lambda_j^{-1}\mathbf{1}_{\lambda_j\neq 0}, j=1,\dots,d$, and set $v := [u_1,\dots,u_d]\overline{\Lambda}w$, then

$$av = [u_1,\dots,u_d]\Lambda[u_1,\dots,u_d]^*[u_1,\dots,u_d]\overline{\Lambda}w = [u_1,\dots,u_d]\Lambda\overline{\Lambda}w = [u_1,\dots,u_d]w$$

since $w = (w_j\mathbf{1}_{\lambda_j\neq 0})_{j=1}^d$. That is, $(t,x)\mapsto v(t,x)\in\mathbb{R}^d$ is a measurable function such that for almost every $t\in[0,T]$ and each i, $a_{i\bullet}(t,x)^*v(t,x) = [a_{ij}\partial_j p_t(x)]$, dx a.e. Finally, $K^p(t,x) := v(t,x)/p_t(x)\mathbf{1}_{p_t(x)>0}$ has the required properties. ∎

Remark 2.6 *The function $v(t,x)$ in the proof of Lemma 2.5 gives an a.e. sense to ∇p_t under $H3)_{P_0}$ as far as we are concerned with the products $a_{i\bullet}^*\nabla p_t$. Clearly, $v(t,x)$ satisfying $a_{i\bullet}(t,x)^*v(t,x) = [a_{ij}\partial_j p_t(x)]$ is not unique a.e. unless $a(t,x)$ is a.e. non singular.*

Under assumptions $H3)_{P_0}$ and $H3)_{Q_0}$ and in view of the previous remark, Lemma 2.5 justifies introducing the following notations:

- $\nabla\ln\frac{p_t}{q_t}(x)$ denotes the equivalence class of the the funcion $K^p(t,x) - K^q(t,x)$ under the relation $f \sim^p g \iff f(t,x) - g(t,x)\in Ker(a(t,x)),\ p_t(x)dx\ dt$ *a.e.*
- $\nabla\frac{p_t}{q_t}(x)$ denotes the equivalence class of the function $\frac{p_t}{q_t}(x)\left(K^p(t,x) - K^q(t,x)\right)$ under the relation $f \sim^q g \iff f(t,x) - g(t,x)\in Ker(a(t,x)),\ q_t(x)dx\ dt$ *a.e.*

It is easily seen that this notation is consistent with the particular case when p_t and q_t are C^1 and strictly positive (i.e. in that case the true gradient belongs to the equivalence class named after it). As customary, we identify equivalence classes with their representatives when the context allows us to do so. Notice then that the relation expected by formal derivation :

$$\frac{p_t}{q_t}(x)\nabla\ln\frac{p_t}{q_t}(x) = \nabla\frac{p_t}{q_t}(x) \tag{2.2}$$

holds true by Lemma 2.5 ii), in the sense that $\frac{p_t}{q_t}(x)k(t,x) \sim^q \frac{p_t}{q_t}(x)\left(K^p(t,x) - K^q(t,x)\right)$ whenever $k(t,x) \sim^p K^p(t,x) - K^q(t,x)$.

Recall now that an element $\mathbb{P}_0 \in \mathcal{M}$ of a given set $\mathcal{M}$ of probability measures in $C([0,T],\mathbb{R}^d)$ is said to be *extremal* if $\mathbb{P}_0 = \alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2$ for some $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}$ and $\alpha \in (0,1)$ implies $\mathbb{P}_0 = \mathbb{P}_1 = \mathbb{P}_2$. We have

Lemma 2.7 *Assume that $H1), H2)_{Q_0}$ and $H3)_{Q_0}$ hold. For each $i = 1, \ldots, d$,*

$$M_t^i := Y_t^i - Y_0^i - \int_0^t \bar{b}_{Q_0}^i(s, Y_s)ds, \ t \in [0,T]$$

is a continuous local martingale w/r to $\mathbb{Q}^{T\to 0}$ and $(\mathcal{G}_t)$, and $\langle M^i, M^j\rangle_t = \int_0^t \bar{a}^{ij}(s,Y_s)ds$ for all $i,j = 1,\ldots,d$. Moreover, if $\mathbb{Q}^{T\to 0}$ is an extremal solution to the martingale problem $(MP)_{Q_0}$, then for any martingale $(N_t)_{t\in[0,T]}$ w/r to $\mathbb{Q}^{T\to 0}$ and $(\mathcal{G}_t)$ such that $N_0 = 0$, there exist predictable processes $(h_t^j)_{t\in[0,T],j=1,\ldots d}$ with $\sum_{i,j=1}^d \int_0^T h_s^j \bar{a}_{ij}(s,Y_s)h_s^i ds < \infty$, $\mathbb{Q}^{T\to 0}$ a.s., and such that $(\int_0^t h_s \cdot dM_s = \sum_{j=1}^d \int_0^t h_s^j dM_s^j)_{t\in[0,T]}$ is a modification of $(N_t)_{t\in[0,T]}$. In particular, $(N_t)_{t\in[0,T]}$ has a continuous modification.

Remark 2.8 *Obviously, $\mathbb{Q}^{T\to 0}$ is an extremal solution to the martingale problem $(MP)_{Q_0}$ if uniqueness holds for it. In particular this is true if pathwise uniqueness for the stochastic differential equation*

$$dX_t = \bar{b}_{Q_0}(t, X_t) + \bar{\sigma}(t, X_t)dW_t, \quad t \in [0,T] \tag{2.3}$$

holds, where $\bar{\sigma}(t,x) = \sigma(T-t,x)$. See Lemma 2.13 below for conditions ensuring pathwise uniqueness which are related to long time stability.

Proof of Lemma 2.7. Standard localization arguments show that M_t^f in $(MP)_{Q_0}$ is a continuous local martingale for any $f \in C^2$ (see e.g. Proposition 2.2 in Ch. VII of [13], its proof for deterministic initial condition also being valid in the general case). Moreover, since $M_t^i = M_t^f$ for $f(x) = x^i$, by Proposition 2.4 , Ch. VII of [13] we get $\langle M^i, M^j\rangle_t = \int_0^t \bar{a}^{ij}(s,Y_s)ds$. The measure $\mathbb{Q}^{T\to 0}$ is therefore a solution to the Problem (12.9) in Jacod [8] in the filtered space $\big(C([0,T],\mathbb{R}^d), (\mathcal{G}_t)_{t\in[0,T]}\big)$, with data given by $\mathcal{G}_0$ and $(Y_t)_{t\in[0,T]}$, and characteristics respectively corresponding to: Q_T as the initial law, the $d-$dimensional process $(B^i = \int_0^\cdot \bar{b}^i(s,Y_s)ds)_{i=1}^d$, the matrix process $(C^{ij} = \int_0^\cdot \bar{a}^{ij}(s,Y_s)ds)_{i,j=1}^d$ and the random measure process on $\mathbb{R}^d$ given by $\mu_t \equiv 0$. The extremality assumption on $\mathbb{Q}^{T\to 0}$ and Theorem 12.21 in [8] imply that any $L^2(\mathbb{Q}^{T\to 0})-$bounded $(\mathcal{G}_t)-$martingale is the sum of on one hand the $L^2(\mathbb{Q}^{T\to 0})$ limit of linear combinations of stochastic integrals with respect to M_t^i,$i = 1,\ldots,d$ and, on the other hand, a compensated jump martingale in the form of stochastic integral with respect to $\mu_t - \nu_t$, with μ_t and ν_t respectively denoting the (trivial) random jump measure associated with the continuous process Y_t and its predictable dual projection (see also Proposition 12.10 in [8]). The statement follows by localization arguments. ∎

We are ready to state the main result of this section. In all the sequel the convention $\inf \emptyset = +\infty$ is adopted. By convenience, we will also assume that the filtration $(\mathcal{G}_t)_{t\in[0,T]}$ is extended to the whole interval $[0,\infty)$ by putting $\mathcal{G}_t = \mathcal{G}_T$ for all $t \in [T,\infty)$.

Theorem 2.9 *Assume that $U : [0,\infty) \to \mathbb{R}$ is a convex function and denote respectively by U'_- and $U''(dy)$ the left-hand derivative of the restriction of U to $(0,+\infty)$ and the non-negative measure on $(0,+\infty)$ equal to the second order distribution derivative of this restriction.*

Let Q_0 and P_0 be probability measures on $\mathbb{R}^d$ such that

$$H_U(P_0|Q_0) < \infty$$

and assume that $H1)$, $H2)_{Q_0}$, $H3)_{Q_0}$ and $H3)_{P_0}$ hold. Suppose moreover that $\mathbb{Q}^{T\to 0}$ is an extremal solution to the martingale problem $(MP)_{Q_0}$. Then, one has

a) *(Stochastic U-entropy dissipation) The submartingale $(U(D_t^T))_{t\in[0,T]}$ has the Doob-Meyer decomposition*

$$\begin{aligned}\forall t \in [0,T],\ U(D_t^T) =&U(D_0^T) + \int_0^t U'_-(D_s^T)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\cdot dM_s \\ &+\frac{1}{2}\int_{(0,+\infty)} L_t^r(D^T)U''(dr) - 1_{\{0<R\le t\}}\Delta U(0),\end{aligned}\tag{2.4}$$

where $R := \inf\{s \in [0,T] : D_s^T = 0\}$, $\Delta U(0) = \lim_{x\to 0^+} U(x) - U(0) \le 0$ and $L_t^r(D^T)$ denotes the local time at level $r \ge 0$ and time t of the continuous version of the martingale $(D_s^T)_{s\in[0,T]}$.

In particular, if U is continuous on $[0,+\infty)$ and C^2 on $(0,+\infty)$, one has

$$\begin{aligned}\forall t \in [0,T],\ U(D_t^T) =&U(D_0^T) + \int_0^t U'(D_s^T)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\cdot dM_s \\ &+\frac{1}{2}\int_0^t U''\left(\frac{p_{T-s}}{q_{T-s}}(Y_s)\right)\left(\nabla^*\left[\frac{p_{T-s}}{q_{T-s}}\right]\bar{a}(s,\cdot)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right]\right)(Y_s)\mathbf{1}_{s<R}ds.\end{aligned}\tag{2.5}$$

b) *(U-Entropy dissipation) We have $\forall t \in [0,T]$,*

$$\begin{aligned}H_U(P_t|Q_t) =& H_U(p_T|q_T) - \Delta U(0)\mathbb{Q}^{T\to 0}(0 < R \le T-t) \\ &+\frac{1}{2}\tilde{\mathbb{E}}^{T\to 0}\left(\int_{(0,+\infty)} L_{T-t}^r(D^T)U''(dr)\right).\end{aligned}\tag{2.6}$$

Last, when U is continuous on $[0,+\infty)$ and C^2 on $(0,+\infty)$,

$$\begin{aligned}H_U(P_0|Q_0) =& H_U(p_T|q_T) \\ &+\frac{1}{2}\int_0^T\int_{\{\frac{p_s}{q_s}(x)>0\}} U''\left(\frac{p_s}{q_s}(x)\right)\left(\nabla^*\left[\frac{p_s}{q_s}\right]a(s,\cdot)\nabla\left[\frac{p_s}{q_s}\right]\right)(x)q_s(x)dxds,\end{aligned}\tag{2.7}$$

where, by a slight abuse of notations, $U''(r)$ denotes the second order derivative of U at point $r > 0$.

Corollary 2.10 *For the choice $U(x) = |x-1|$, under the assumptions of Theorem 2.9 and in particular if $P_0 \ll Q_0$, the total variation distance $\|P_t - Q_t\|_{\mathrm{TV}}$ satisfies*

$$\forall t \le T,\ \|P_t - Q_t\|_{\mathrm{TV}} = \|P_0 - Q_0\|_{\mathrm{TV}} - \tilde{\mathbb{E}}^{T\to 0}(L_T^1(D^T) - L_{T-t}^1(D^T)).$$

When $\frac{p_t}{q_t}(Y_t)$ is a continuous $\mathbb{Q}^T$-$\mathcal{G}_t$ semimartingale and in particular if $(t,x)\mapsto \frac{p_t}{q_t}(x)$ is well-defined and of class $C^{1,2}$, we deduce that

$$\forall t \le T,\ \|P_t - Q_t\|_{\mathrm{TV}} = \|P_0 - Q_0\|_{\mathrm{TV}} - \tilde{\mathbb{E}}^T(L_t^1(\frac{p_\cdot}{q_\cdot}(Y_\cdot))).$$

Remark 2.11 *a) Notice when $H_U(P_0|Q_0) < \infty$ that $H2)_{Q_0}$ implies $H2)_{P_0}$ by Lemma 2.1.*

b) If condition $C)_{Q_0}$ in Remark 2.4 i) holds (and thus $H3)_{Q_0}$ holds), then also $C)_{P_0}$ (and thus $H3)_{P_0}$) holds if for instance $\frac{dP_0}{dQ_0}$ has at most polynomial growth.

To prove Theorem 2.9 we first obtain explicit expressions for the martingale $(D_t^T)_{t\in[0,T]}$ introduced in Lemma 2.1, relying on the extremality assumption in order to use Girsanov theory in the absolutely-continuous setting. The last of the three following assertions will not be needed but provides additional information about that process.

Lemma 2.12 *Assume that $H1)$, $H2)_{Q_0}$, $H3)_{Q_0}$ and $H3)_{P_0}$ hold together. Suppose moreover that $P_0 \ll Q_0$ and that $\mathbb{Q}^{T\to 0}$ is an extremal solution to the martingale problem $(MP)_{Q_0}$. Let $(D_t^T)_{t\in[0,T)}$ be the Girsanov density process defined in Lemma 2.1.*

i) With R the $(\mathcal{G}_t)$-stopping time $R := \inf\{s \in [0,T] : D_s^T = 0\}$, we have $\mathbb{Q}^{T\to 0}-$a.s. that

$$\forall t \in [0,T], \ \int_0^t \left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\, ds < \infty, \ \text{and}$$
$$\forall t \in [0,R), \ \int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds < \infty \ \text{on } \{R>0\}.$$

ii) The process $(D_t^T)_{t\in[0,T]}$ has a continuous version, denoted in the same way, such that

$$\begin{aligned}\mathbb{Q}^{T\to 0}\ a.s,\ \forall t \in [0,T],\ D_t^T =& \frac{p_T}{q_T}(Y_0) + \int_0^t \nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\cdot dM_s \\ =& \frac{p_T}{q_T}(Y_0) + \int_0^t \nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{\{\frac{p_{T-s}}{q_{T-s}}(Y_s)>0\}}\cdot dM_s \\ \text{and } \langle D^T\rangle_t =& \int_0^t \left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\, ds.\end{aligned}$$

iii) Finally, if we define the $(\mathcal{G}_t)$-stopping times $\tau^o := 0\cdot\mathbf{1}_{D_0^T=0} + \infty\cdot\mathbf{1}_{D_0^T>0}$ and

$$\tau := \inf\left\{t\in[0,T] : \int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds = \infty\right\},$$

then $\mathbb{Q}^{T\to 0}-$a.s. $R = \tau\wedge\tau^o$, and $\forall t\in[0,T]$,

$$\begin{aligned}&D_t^T = \mathbf{1}_{\{t<\tau\}}\frac{dp_T}{dq_T}(Y_0)\times \\ &\exp\left\{\int_0^t \nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\cdot dM_s - \frac{1}{2}\int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds\right\}.\end{aligned} \tag{2.8}$$

In particular, on $\{R>0\}$ the second integral in i) is a.s. divergent at $t=R$.

Proof of Lemma 2.12. By Lemma 2.7, the $\mathbb{Q}^{T\to 0}$-martingale $(D_t^T)_{t\in[0,T]}$ admits the continuous version $D_0^T + \sum_{j=1}^d \int_0^t h_s^j dM_s^j$ still denoted by D_t^T for simplicity. The martingale

representation property and standard properties of stochastic integrals moreover imply that D_t^T is determined by the processes $\langle D^T, M^i\rangle = \int_0^\cdot \sum_{j=1}^d h_t^j \bar{a}_{ij}(t,Y_t)dt$, $i=1,\dots,d$. Consequently, h_t can be replaced (leaving D_t^T unchanged) by any predictable process k_t such that for each i, $\int_0^\cdot \sum_{j=1}^d h_t^j \bar{a}_{ij}(t,Y_t)dt = \int_0^\cdot \sum_{j=1}^d k_t^j \bar{a}_{ij}(t,Y_t)dt$ $\mathbb{Q}^{T\to 0}$ a.s. ($\int_0^T \sum_{i,j=1}^d k_s^j \bar{a}_{ij}(s,Y_s)k_s^i ds = \int_0^T \sum_{i,j=1}^d h_s^j \bar{a}_{ij}(s,Y_s)h_s^i ds < \infty$ $\mathbb{Q}^{T\to 0}$ a.s. immediately follows). Furthermore, since $D_t^T = D_{t\wedge R}^T$, we may and shall assume that $\mathbb{Q}^{T\to 0}$ a.s. $h_t = h_t \mathbf{1}_{t<R} = h_t \mathbf{1}_{D_t^T>0}$ for all $t\in[0,T]$. Let us also notice that, by Fubini's Theorem, it $\mathbb{Q}^{T\to 0}-$a.s. holds that $D_s^T(Y_s) = \frac{p_{T-s}}{q_{T-s}}(Y_s)$ (and then $\mathbf{1}_{\{R>s\}} = \mathbf{1}_{\{\frac{p_{T-s}}{q_{T-s}}(Y_s)>0\}}$) for a.e. $s\in[0,T]$.

Now, by our assumptions and Theorem 2.3 a), $\mathbb{P}^{T\to 0} \ll \mathbb{Q}^{T\to 0}$ are probability measures respectively solving the martingale problems $(MP)_{P_0}$ and $(MP)_{Q_0}$. The processes $\int_0^\cdot \bar{b}_{P_0}^i(t,Y_t)dt$ and $\int_0^\cdot \bar{b}_{Q_0}^i(t,Y_t)dt + \int_0^\cdot (D_t^T)^{-1} h_t^j d\langle M^i, M^j\rangle_t$ then are $\mathbb{P}^{T\to 0}-$ indistinguishable (see e.g. Proposition 12.18 v) in [8]). Using these facts, the expression for $\langle M^i, M^j\rangle$ in Lemma 2.7 and part ii) of Lemma 2.5 we deduce first that, $\mathbb{P}^{T\to 0}-$a.s.,

$$\bar{b}_{P_0}^i(t,Y_t) - \bar{b}_{Q_0}^i(t,Y_t) = \sum_{j=1}^d \bar{a}_{ij}(t,Y_t)^* \left(h_t^j \frac{q_{T-t}}{p_{T-t}}(Y_t)\right) = \bar{a}_{i\bullet}(t,Y_t)^*(K^p(T-t,Y_t) - K^q(T-t,Y_t)) \tag{2.9}$$

for a.e. $t\in[0,T]$ and each i. By part ii) of Lemma 2.5 we then also get

$$\int_0^\cdot \sum_{j=1}^d h_t^j \bar{a}_{ij}(t,Y_t)dt = \int_0^\cdot \bar{a}_{i\bullet}(t,Y_t)^*(K^p(T-t,Y_t) - K^q(T-t,Y_t))\frac{p_{T-t}(Y_t)}{q_{T-t}(Y_t)}dt,\ i=1,\dots,d,$$

$\mathbb{P}^{T\to 0}-$a.s., and then $\mathbb{Q}^{T\to 0}-$a.s. because of our assumption on h. From these identities and our previous discussion we deduce the that we can choose $h_t = \nabla \frac{p_{T-t}}{q_{T-t}}(Y_t)\mathbf{1}_{\{\frac{p_{T-t}}{q_{T-t}}(Y_t)>0\}} = \nabla \frac{p_{T-t}}{q_{T-t}}(Y_t)\mathbf{1}_{\{R>t\}}$. This proves part ii). The first property of the process $\nabla \frac{p_{T-t}}{q_{T-t}}(Y_t)$ in i) is thus consequence of the general properties of h in the representation formula for D_t^T. The second assertion in i) easily follows from the first one, taking into account the definitions of $\nabla \frac{p_{T-t}}{q_{T-t}}(Y_t)$ and $\nabla \ln \frac{p_{T-t}}{q_{T-t}}(Y_t)$, the relation (2.2) (in its rigorous sense) and the properties of D_t^T.

To establish iii), we again use the extremality of $\mathbb{Q}^{T\to 0}$ in order to apply Theorem 12.48 in [8]. Thanks to part ii) of Lemma 2.5 and equation (2.9), the objects z, K, B and T_n in (12.32), (12.35) and (12.42) of [8] alluded in that theorem, correspond in our setting to, respectively, $\frac{p_T}{q_T}(Y_0)$, $\nabla \ln\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)$, the increasing process $A_t := \int_0^t \left(\nabla\left[\ln \frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln \frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds$ and the stopping time $\tau_n := \inf\{t\in[0,T] : A_t \geq n\}$. This and Lemma 12.36 d) in [8], yield the fact that D_t^T equals (2.8), $\mathbb{Q}^{T\to 0}$ a.s in the set $\cup_{n\in\mathbb{N}}\{t\in[0,T] : t\leq \tau_n\}$. It is also established therein that $\tau_n \nearrow \tau$ $\mathbb{Q}^{T\to 0}$ a.s., impling that the latter random set equals the interval $[0,\tau)\cap[0,T]$. Moreover, on this interval, the integrals which appear in the exponential factor in (2.8) are finite. Therefore, either $\tau^o = 0$ and then $R=0$, or $\tau^o=\infty$ and then $R\geq\tau$.

By Theorem 12.48 in [8] as well, we have $D_t^T = \liminf_{n\to\infty} D_{\tau_n}^T$ for t in $[\tau,T]$, $\mathbb{Q}^{T\to 0}$ a.s. Thus, $t\mapsto D_t^T$ is constant in $[\tau,T]$, $\mathbb{Q}^{T\to 0}$ a.s.. By Theorem 12.39 in [8] we have $\mathbb{P}^{T\to 0}(\tau<\infty)=0$. Since $\mathbb{P}^{T\to 0}(\tau<\infty) = \mathbb{P}^{T\to 0}(\tau\leq T) = \mathbb{Q}^{T\to 0}(1_{\{\tau\leq T\}}D_T^T)$, the a.s. constancy of $t\mapsto D_t^T$ on $[\tau,T]$ ensures that $\mathbb{Q}^{T\to 0}$ a.s. $D_t^T=0$ for all $t\in[\tau,T]$, when the latter interval is non empty. As a consequence $\mathbb{Q}^{T\to 0}$ a.s., $R\leq\tau$ so that $R=\tau\wedge\tau^o$. This completes the proof. ∎

Proof of Theorem 2.9. Since $(D_t^T)_{t\in[0,T]}$ is a continuous non-negative $\mathbb{Q}^{T\to 0}$-martingale

and U'_- is locally bounded on $(0,+\infty)$, $t \mapsto \int_0^t \left[U'_-(D_s^T)\right]^2 d\langle D^T\rangle_s$ is finite and continuous on $[0,T]$ when $R > T$ and finite and continuous on $[0,R)$ otherwise. In the latter case, $\int_0^R \left[U'_-(D_s^T)\right]^2 d\langle D^T\rangle_s$ makes sense but is possibly infinite. Define for any positive integer n the stopping time

$$R_n := \inf\left\{t \in [0, T \wedge R] : D_t^T \leq \frac{1}{n} \text{ or } \int_0^t \left[U'_-(D_s^T)\right]^2 d\langle D^T\rangle_s \geq n\right\}.$$

For all $t \in [0,T]$, $\int_0^{t\wedge R_n} \left[U'_-(D_s^T)\right]^2 d\langle D^T\rangle_s \leq n$ and $\mathbb{E}\left(\int_0^{t\wedge R_n} U'_-(D_s^T) dD_s^T\right) = 0$. Moreover $R_n \nearrow R$ as $n \to \infty$.

Let $t \in [0,T]$. By Tanaka's formula,

$$U(D_{t\wedge R_n}^T) = U(D_0^T) + \int_0^{t\wedge R_n} U'_-(D_s^T) dD_s^T + \frac{1}{2}\int_{(0,+\infty)} L_{t\wedge R_n}^r(D^T) U''(dr). \tag{2.10}$$

The assumption that $H_U(P_0|Q_0) < \infty$ and Remark 2.2 a) imply that $(U(D_s^T))_{s\in[0,T]}$ is a uniformly integrable $\mathbb{Q}^{T\to 0}$-submartingale. Since the $\mathbb{Q}^{T\to 0}$-expectation of the stochastic integral is zero, one deduces

$$\tilde{\mathbb{E}}^{T\to 0}\left(U(D_{t\wedge R_n}^T)\right) = \tilde{\mathbb{E}}^{T\to 0}(U(D_0^T)) + \frac{1}{2}\tilde{\mathbb{E}}^{T\to 0}\left(\int_{(0,+\infty)} L_{t\wedge R_n}^r(D^T) U''(dr)\right).$$

When $n \to \infty$, since U is continuous on $(0,+\infty)$ by convexity, $U(D_{t\wedge R_n}^T)$ converges to $U(D_{t\wedge R}^T) + \Delta U(0) 1_{\{0<R\leq t\}} = U(D_t^T) + \Delta U(0) 1_{\{0<R\leq t\}}$ and by uniform integrability, $\mathbb{E}(U(D_{t\wedge R_n}^T))$ converges to $\mathbb{E}(U(D_t^T)) + \Delta U(0)\mathbb{Q}^{T\to 0}(0 < R \leq t)$. Dealing with the expectation of the integral in the right-hand-side by monotone convergence, one obtains

$$\mathbb{E}(U(D_t^T)) = \tilde{\mathbb{E}}^{T\to 0}(U(D_0^T)) - \Delta U(0)\mathbb{Q}^{T\to 0}(0 < R \leq t) + \frac{1}{2}\tilde{\mathbb{E}}^{T\to 0}\left(\int_{(0,+\infty)} L_{t\wedge R}^r(D^T) U''(dr)\right).$$

Since according to Lemma 2.12 ii), D^T is equal to zero on $[R,T]$, one can replace $t \wedge R$ by t in the last expectation. Replacing t by $T-t$ in this equation, one gets (2.6). Moreover $\mathbb{Q}^{T\to 0}$ a.s., $\int_{(0,+\infty)} L_t^r(D^T) U''(dr)$ is the finite limit of the integral with respect to $U''(dr)$ in the right-hand-side of (2.10) as $n \to \infty$. Since the left-hand-side converges to $U(D_t^T) + \Delta U(0) 1_{\{0<R\leq t\}}$ we deduce that the stochastic integral in the right-hand-side also has a finite limit. Hence $\int_0^{t\wedge R} \left[U'(D_s^T)\right]^2 d\langle D^T\rangle_s < +\infty$, $\int_0^{t\wedge R} U'(D_s^T) dD_s^T$ makes sense and (2.4) holds. When U is continuous on $[0,+\infty)$ and C^2 on $(0,+\infty)$, (2.6) written for $t = 0$ combined with the occupation times formula and Lemma 2.12 imply (2.5) and that

$$\begin{aligned} H_U(P_0|Q_0) =& H_U(p_T|q_T) \\ &+ \frac{1}{2}\tilde{\mathbb{E}}^{T\to 0}\left(\int_0^T U''(D_s^T) 1_{\{s<R\}} \left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s) ds\right). \end{aligned}$$

Since Y_s admits the density q_{T-s} and for almost all $s \in [0,T)$, $D_s^T = \frac{p_{T-s}}{q_{T-s}}(Y_s)$ and $\{R > s\} = \{\frac{p_{T-s}}{q_{T-s}}(Y_s) > 0\}$, by changing variables $s \mapsto T - s$ we have established (2.7). ∎

We end this section with the following two statements concerning the important case when $Q_0(dx) = q_0(x)dx = p_\infty(x)dx$ is a stationary probability law for the Markov diffusion (0.2).

Lemma 2.13 *Suppose that the functions σ and b do not depend on time and that the Markov diffusion process* (0.1) *has an invariant density $p_\infty(x)$, such that $\int_D |\partial_j (a_{ij}(x)p_\infty(x))|\,dx < \infty$ for any open bounded set $D \subset \mathbb{R}^d$, where $a = \sigma\sigma^*$ and the derivatives are meant in the distribution sense. Set for $i = 1, \ldots, d$*

$$\Psi^i(x) := -\sum_{j=1}^{d'} \frac{\partial_j (a_{ij}(x)p_\infty(x))}{p_\infty(x)} \ \textit{if } p_\infty(x) > 0$$

and 0 otherwise, and $\Psi = (\Psi^i)_{i=1}^d$. Finally, assume that $Q_0(dx) = p_\infty(x)dx$, that $H1)$ and $H2)_{Q_0}$ hold , and that Ψ is the sum of a locally Lipschitz continuous function $\hat{\Psi}$ and a monotone function $\tilde{\Psi}$:

$$\langle \tilde{\Psi}(x) - \tilde{\Psi}(y), x - y \rangle \geq 0 \ \textit{ for all } x, y \in \mathbb{R}^d.$$

Then, pathwise uniqueness holds for the stochastic differential equation (3.1)*. In particular, $\mathbb{Q}^{T\to 0}$ is an extremal solution to the martingale problem $(MP)_{Q_0}$.*

Observe that if $d = d'$, $\sigma = Id$ is the identity matrix, and $p_\infty(x) = Ce^{-2V(x)}$ for some convex function $V : \mathbb{R}^d \to \mathbb{R}$, then $\Psi(x) = 2\nabla V(x)$ satisfies the monotonicity condition. More generally, if the matrix a has locally Lipschitz derivates, then Ψ satisfies the above condition for instance if $x \mapsto a(x)K^{p_\infty}(x)$ is moreover monotone or, alternatively, if $p_\infty(x)$ is strictly positive and has locally bounded second oder derivatives.

Proof . Let X_t and Y_t be two solutions to (3.1) constructed on the same probability space and equal at $t = 0$. By Itô's formula and the assumption on $\tilde{\Psi}$ we get

$$\begin{aligned}
|X_t - Y_t|^2 \leq{}& 2\int_0^t (X_s^i - Y_s^i)\left(\sigma^{ij}(X_s) - \sigma^{ij}(Y_s)\right) dW_s^j \\
&- 2\int_0^t (X_s^i - Y_s^i)\left(b^i(X_s) - b^i(Y_s) + \hat{\Psi}^i(X_s) - \hat{\Psi}^i(Y_s)\right) ds \\
&+ \int_0^t tr(\sigma(X_s) - \sigma(Y_s))(\sigma(X_s) - \sigma(Y_s))^* ds.
\end{aligned}$$

Thanks to the local Lipschitz-continuity of b, σ and $\hat{\Psi}$, and after localizing, taking expectations, and using the BDG inequality, it is standard to conclude with Gronwall's lemma that $X = Y$. ∎

Proposition 2.14 *Assume that the functions σ and b do not depend on time and are Lipschitz continuous. Assume moreover that the Markov diffusion process* (0.2) *has an invariant density $p_\infty(x)$ and a strictly positive transition density $\varphi_t(x, y)$ w.r.t. dy, which is continuous in (x, y) for each $t > 0$. Last, assume that $H_U(P_t|Q_t) < \infty$ for some $t \geq 0$. Then*

$$\lim_{s\to\infty} H_U(P_s|Q_s) = 0.$$

Remark 2.15 *For conditions ensuring the joint continuity of the transition density with respect to (x, y), we refer to [5] Chapter 9 under uniform ellipticity and to [10] Theorem 4.5 under hypoellipticity.*

Proof . According to Corollary 1.2, it is enough to check that the tail σ-field $\cap_{s\geq 0}\sigma(X_r^{Q_0}, r \geq s)$ is trivial a.s..
First, by our assumptions on the coefficients σ and b, the semigroup $(P_t)_{t\geq 0}$ associated with (0.2) is Feller, and moreover strongly Feller by the continuity in (x, y) of $\varphi_t(x, y)$ (it is enough that $P_t f$ be continuous for all f such that $0 \leq f \leq 1$, which is true because $P_t f$ and $P_t(1-f)$ are both l.s.c. functions summing 1, by Fatou's Lemma).

The positivity of the transition density implies that any invariant probability measure is equivalent to the Lebesgue measure on $\mathbb{R}^d$. Therefore $p_\infty(x)dx$ is the unique invariant measure, which thus is ergodic. Moreover, $p_\infty(x) > 0$ dx a.e.. Let $\mathbb{P}_\infty$ denote the law of the solution to (0.2) starting from an initial condition distributed according to $p_\infty(x)dx$ and $(Y_t)_{t\geq 0}$ denote the canonical process on $C([0, +\infty), \mathbb{R}^d)$. For simplicity we also write $\mathbb{P}_x$ instead of $\mathbb{P}_{0,x}$. By the ergodic theorem and the strict positivity of p_∞, we have $\int_0^\infty \mathbf{1}_A(Y_t)dt = +\infty$, $\mathbb{P}_\infty-$ a.s. for each Borel set A in $\mathbb{R}^d$ with strictly positive Lebesgue measure. If $\tilde{A} = \{y \in C([0, +\infty), \mathbb{R}^d) : \int_0^\infty \mathbf{1}_A(y_s)ds = \infty\}$, we deduce that $\mathbb{P}_x(\tilde{A}) = 1$ for $dx-$ almost every x. But $\tilde{A}$ is a tail event, and by the Markov property one has $\mathbb{P}_x(\tilde{A}) = \mathbb{E}_x(\mathbb{P}_{Y_t}(\tilde{A}^t))$ for any $t > 0$ and a suitable measurable set $\tilde{A}^t$ of $C([0, +\infty), \mathbb{R}^d)$. The strong Feller property then yields $\mathbb{P}_x(\tilde{A}) = 1$ for all $x \in \mathbb{R}^d$. That is, X defined by (0.2) is Harris recurrent.

By Theorem 1.3.9 in [9] (and noting that its proof uses only continuity of $\varphi_t(x, y)$ in (x, y) for each $t > 0$ but not continuity in (t, x, y)), we get that any tail event B is such that $\mathbb{P}_x(B) = 1$ for all $x \in \mathbb{R}^d$ or $\mathbb{P}_x(B) = 0$ for all $x \in \mathbb{R}^d$, which concludes the proof. ∎

Remark 2.16 *Under the positivity assumption on the transition density, P_t and Q_t admit positive densities p_t and q_t as soon as $t > 0$. For the choice $U(x) = |x-1|$, $H_U(P_t|Q_t) = \int_{\mathbb{R}^d} |p_t(x) - q_t(x)|dx$ is equal to the total variation distance between P_t and Q_t. According to Theorem 1.3.8 [9], the tail σ-field is trivial a.s. if and only if this total variation distance converges to 0 for all choices of the initial distributions P_0 and Q_0.*

3 Dissipation of the Fisher information and non-intrisic Bakry Emery criterion

We will from now on focus in the case when $Q_0(dx) = p_\infty(x)dx$ is a stationary probability law for the Markov diffusion (0.1) . We denote

$$I_U(p_s|p_\infty) = \frac{1}{2}\int_{\{\frac{p_s}{p_\infty}>0\}} U''\left(\frac{p_s}{p_\infty}\right)\left(\nabla^*\left[\frac{p_s}{p_\infty}\right] a\nabla\left[\frac{p_s}{p_\infty}\right]\right) p_\infty dx$$

the integral that appears in the right-hand-side of (2.7), and we refer to it as the $U-$ Fisher information.

Inspired by the famous Bakry-Emery approach, we want to compute the derivative of $I_U(p_s|p_\infty)$ with respect to the time variable.

In all the sequel, we make the following assumptions :

$H4)$ The drift function b is time-homogeneous and has first order derivatives which are globally bounded and Hölder-continuous uniformly in $\mathbb{R}^d$, and

the matrix σ is time-homogeneous and has up to second order derivatives which are globally bounded and Hölder-continuous uniformly in $\mathbb{R}^d$.

$H5)_{p_\infty}$ The Markov process defined by (0.1) has an invariant density $p_\infty(x)$, and $Q_0(dx) = p_\infty(x)dx$. Moreover, p_∞ has locally bounded derivatives up to the second order which are globally Hölder continuous, and $p_\infty(x) > 0$ for all $x \in \mathbb{R}^d$.

$H6)_{p_0}$ The initial distribution P_0 admits a probability density p_0 with respect to the Lebesgue measure. Moreover, we assume that $H3)_{p_0}$ holds, and that $p_t(x) = \frac{dP_t}{dx}(x)$ has space derivatives up to the second order for each $t > 0$, which are continuous in $(t, x) \in (0, T] \times \mathbb{R}^d$ and bounded and Hölder continuous in $x \in \mathbb{R}^d$ uniformly in $(t, x) \in [\delta, T] \times \mathbb{R}^d$ for each $\delta \in (0, T]$.

Notice that $H4)$ implies $H1)$. Next, $H5)_{p_\infty}$ implies $H2)_{Q_0}$ for $Q_0(dx) = p_\infty(x)dx$ and combined with $H1)$ (or with $H4)$), it implies $H3)_{Q_0}$. Assumptions $H4)$ and $H5)_\infty$ together imply by Lemma 2.13 that uniqueness holds for the martingale problem $(MP)_{p_\infty}$. Therefore the hypotheses of Theorem 2.9 hold within the present Section. If $H5)_\infty$ and $H6)_{p_0}$ hold, $\nabla \frac{p_t}{p_\infty}$ is defined everywhere. Since the gradient of a C^1 non-negative function is equal to 0 when this function is equal to 0, $\nabla \frac{p_t}{p_\infty}$ is equal to $\frac{p_t}{p_\infty}\left(\mathbf{1}_{p_t>0}\frac{\nabla p_t}{p_t} - \frac{\nabla p_\infty}{p_\infty}\right)$ and belongs to the equivalence class (with respect to p_∞) defined after Remark 2.6. We will throughout in the sequel use this everywhere defined representative, in particular in Equation (2.7) which states that the U-entropy dissipation is equal to the U-Fisher information.

Under $H4)$, if moreover a and b are bounded with a uniformly elliptic, then $H6)_{p_0}$ holds for any compactly supported probability density p_0, by [5] Chapter 9. We refer to [10] for conditions ensuring that $H6)_{p_0}$ holds under hypoellipticity.

Let us establish some notation.

- We write $\mathbb{P}_\infty^{T\to 0} := \mathbb{Q}^{T\to 0}$ and $\bar{b}_i := \bar{b}^i_{Q_0}$, $i = 1, \dots, d$.
- $(A^{-1})_{kl}$ denotes the (k, l) coordinate of the inverse A^{-1} of an invertible matrix A.
- By possibly enlarging the probability space $\mathcal{G}_t - \mathbb{P}_\infty^{T\to 0}$, we introduce a Brownian motion $\bar{W}$ such that Y_t solves the stochastic differential equation :

$$dY_t = \bar{b}(Y_t)dt + \sigma(Y_t)d\bar{W}_t, \quad t \in [0, T] \text{ where } \bar{b}_i(y) = -b(y) + \frac{\partial_j(a_{ij}(y)p_\infty(y))}{p_\infty(y)}. \tag{3.1}$$

  By Lemma 2.13, under assumptions $H4)$ and $H5)_\infty$, existence of a unique strong solution holds for this SDE.
- We write $\rho_t(x) := \frac{p_{T-t}}{p_\infty}(x)$, $t \in [0, T]$.

We will make use of the stochastic flow defined by the two-parameter process $\xi_t(x)$ satisfying

$$d\xi^i_t(x) = \sigma_{ik}(\xi_t(x))d\bar{W}^k_t + \bar{b}_i(\xi_t(x))dt, \quad (t, x) \in [0, T) \times \mathbb{R}^d,\ i = 1, \dots d, \tag{3.2}$$

and $\xi_0(x) = x$, and we notice that $\xi_t(Y_0) = Y_t$. We shall also deal with the family of continuous $\mathcal{G}_t - \mathbb{P}_\infty^{T\to 0}-$ local martingales $(D_t(x) = \rho_t(\xi_t(x)) : t \in [0, T])_{x\in\mathbb{R}^d}$ defined by

$$dD_t(x) = [\sigma_{ik}\partial_i\rho]\,(t, \xi_t(x))d\bar{W}^k_t \quad , \quad D_0(x) = \frac{p_T}{p_\infty}(x) = \rho_0(x). \tag{3.3}$$

According to Lemmas 2.7 and 2.12 and Equation (3.1), $D_t(Y_0)$ is equal to the process D_t^T considered in the previous section. Writing $\nabla\rho_t(\xi_t(x)) = (\nabla_x\xi_t(x))^{-1}\nabla_x[\rho_t(\xi_t(x))]$, we remark that $d\nabla\rho_t(\xi_t(x))$ can be obtained with the Itô product rule, by computing $d(\nabla_x\xi_t(x))^{-1}$ and $d\nabla_x[\rho_t(\xi_t(x))]$, as we do in the two next Lemmas:

Lemma 3.1 *The process $(t,x)\mapsto\xi_t(x)$ has a $\mathbb{P}_\infty^{T\to 0}$ a.s. continuous version such that the mapping $x\mapsto\xi_t(x)$ is a global diffeomorphism of class $C^{1,\alpha}$ for some $\alpha\in(0,1)$ and every $t\in[0,T]$. Moreover, we have*

$$d\partial_j\xi_t^i(x) = \partial_p\sigma_{ik}(t,\xi_t(x))\partial_j\xi_t^p(x)d\bar{W}_t^k + \ \partial_p\bar{b}_i(t,\xi_t(x))\partial_j\xi_t^p(x)dt, \quad (t,x)\in[0,T)\times\mathbb{R}^d \tag{3.4}$$

with $\partial_j\xi_0^i(x)=\delta_{ij}$. Finally, writing $\nabla\xi_t(x) = (\partial_j\xi_t^i(x))_{ij}$, it holds that

$$\begin{aligned} d(\nabla\xi_t(x))^{-1}_{kl} =& -(\nabla\xi_t(x))^{-1}_{ki}[\partial_l\sigma_{ir}](t,\xi_t(x))d\bar{W}_t^r - \nabla\xi_t(x))^{-1}_{ki}[\partial_l\bar{b}_i](t,\xi_t(x))dt \\ &+(\nabla\xi_t(x))^{-1}_{ki}[\partial_m\sigma_{ir}\partial_l\sigma_{mr}](t,\xi_t(x))dt, \qquad (t,x)\in[0,T)\times\mathbb{R}^d. \end{aligned} \tag{3.5}$$

Proof . Under assumptions $H4)$ and $H5)_\infty$, classic results by Kunita [9] (see Theorem 4.7.2) imply the asserted regularity properties of the stochastic flow, as well as the $\mathbb{P}_\infty^{T\to 0}$ a.s. existence of the inverse matrix $(\nabla\xi_t(x))^{-1}$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$. Since the smooth map $A\mapsto A^{-1}$, defined on non singular matrices, has first and second derivatives respectively given by the linear and bilinear operators $F\mapsto -A^{-1}FA^{-1}$ and $(F,K)\mapsto A^{-1}FA^{-1}KA^{-1}+A^{-1}KA^{-1}FA^{-1}$ (where F,K are generic square-matrices), we deduce that for $A=(a_{ij})_{i,j=1\dots d}$,

$$\frac{\partial(A^{-1})_{kl}}{\partial a_{ij}} = -A^{-1}_{ki}A^{-1}_{jl}, \quad \text{and} \quad \frac{\partial^2(A^{-1})_{kl}}{\partial a_{ij}\partial a_{mn}} = A^{-1}_{ki}A^{-1}_{jm}A^{-1}_{nl} + A^{-1}_{km}A^{-1}_{ni}A^{-1}_{jl}$$

for all $k,l,i,j,m,n\in\{1,\dots,d\}$. Equation 3.5 follows by applying Itô's formula to each of the functions $A\mapsto(A^{-1})_{kl}$ and the semimartingales $(\partial_j\xi_t^i(x))$, $i,j=1\dots d$.

∎

Lemma 3.2 *The process $D_t(x)$ has a modification still denoted $D_t(x)$ such that $\mathbb{P}_\infty^{T\to 0}$ a.s. the function $(t,x)\mapsto D_t(x)$ is continuous and $x\mapsto D_t(x)$ is of class C^1 for each t. This modification is indistinguishable from $(\rho_t(\xi_t(x)):(t,x)\in[0,T)\times\mathbb{R}^d)$ and we have*

$$d\partial_k D_t(x) = \partial_m\left[\sigma_{ir}\partial_i\rho\right](t,\xi_t(x))\partial_k\xi_t^m(x)d\bar{W}_t^r = d\left[\partial_m\rho(t,\xi_t(x))\partial_k\xi_t^m(x)\right] \tag{3.6}$$

for all $(t,x)\in[0,T)\times\mathbb{R}^d$.

Proof . Thanks to assumption $H6)_{p_0}$ and the regularity of $x\mapsto\xi_t(x)$ established in Lemma 3.1, the statements follow from Theorem 3.3.3 of Kunita [9] (see also Exercise 3.1.5 therein). ∎

Evaluating expressions (3.5) and (3.6) in $x=Y_0$, we obtain using Itô's product rule that

$$\begin{aligned} d\partial_l\rho_t(Y_t) &= \left[\sigma_{kr}\partial_{lk}\rho\right](t,Y_t)d\bar{W}_t^r - \left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr} + \partial_k\rho\partial_l\bar{b}_k\right](t,Y_t)dt \\ &= \left[\sigma_{kr}\partial_{lk}\rho\right](t,Y_t)d\bar{W}_t^r - \left[\frac{1}{2}\partial_{kj}\rho\partial_l a_{kj} + \partial_k\rho\partial_l\bar{b}_k\right](t,Y_t)dt \end{aligned} \tag{3.7}$$

From now on, for notational simplicity the argument (t,Y_t) will sometimes be omitted.

To compute the dissipation of the U-Fischer information, in all the sequel we make the following regularity assumption on U:

$H7)$ The convex function $U:[0,\infty)\to\mathbb{R}$ is of class C^4 on $(0,+\infty)$, continuous on $[0,+\infty)$ and satisfies $U(1)=U'(1)=0$.

The assumption that $U'(1)=0$ is inspired in the analysis on admissible entropies developed in Arnold et al. [1]. It is granted without modifying the functions $p\mapsto H_U(p|p_\infty)$ and $p\mapsto I_U(p|p_\infty)$ by replacing $U(r)$ by $U(r)-U'(1)(r-1)$ if needed. Notice that if $H7)$ holds, $U(r)$ attains the minimum 0 at $r=1$ and therefore $U\geq 0$ by convexity.

We do not assume that the entropy function U is C^4 on the closed interval $[0,+\infty)$, since we want to deal with $U(r)=r\ln(r)-(r-1)$. That is why we introduce some regularization U_δ indexed by a positive parameter δ : we chose U_δ such that $U_\delta(r)=U(r+\delta)$ for $r\geq 0$ and U_δ is extended to a C^4 function on $\mathbb{R}$.

Proposition 3.3 *One has*

$$d\left[U''_\delta(\rho)\nabla^*\rho a\nabla\rho\right]=tr(\Lambda_\delta\Gamma)dt+U''_\delta(\rho)\bar{\theta}dt+d\hat{M}^{(\delta)}$$

where $\hat{M}^{(\delta}$ *is the* $\mathcal{G}_t-\mathbb{P}_\infty^{T\to 0}-$*local martingale*

$$d\hat{M}^{(\delta)}:=\left\{2\;U''_\delta(\rho)\sigma_{l'i}\;\partial_{l'}\rho\partial_k\left[\partial_l\rho\sigma_{li}\right]+\left[\nabla\rho^*a\nabla\rho\right]U^{(3)}_\delta(\rho)\partial_k\rho\right\}\sigma_{kr}d\bar{W}^r=\partial_k\left[U''_\delta(\rho)\nabla^*\rho a\nabla\rho\right]\sigma_{kr}d\bar{W}^r,$$

Λ_δ *and* Γ *are the square matrices defined by*

$$\Lambda_\delta:=\begin{bmatrix}U''_\delta(\rho) & U^{(3)}_\delta(\rho)\\ U^{(3)}_\delta(\rho) & \frac{1}{2}U^{(4)}_\delta(\rho)\end{bmatrix}\qquad\Gamma:=\begin{bmatrix}\nabla^*(\sigma_{\bullet i}\cdot\nabla\rho)a\nabla(\sigma_{\bullet i}\cdot\nabla\rho) & (\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\;a\nabla(\sigma_{\bullet i}\cdot\nabla\rho)\\ (\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\;a\nabla(\sigma_{\bullet i}\cdot\nabla\rho) & |\nabla^*\rho a\nabla\rho|^2\end{bmatrix}$$

and

$$\begin{aligned}\bar{\theta}=2\Big\{&\left[\sigma_{l'i}\partial_{l'}\rho a_{mk}\partial_m\sigma_{li}\partial_{lk}\rho\right]+\sigma_{l'i}\partial_{l'}\rho\partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li}+\frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right]\\&-a_{ll'}\partial_{l'}\rho\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr}+\partial_k\rho\partial_l\bar{b}_k\right]\Big\}\\=2\Big\{&\partial_{l'}\rho\partial_l\rho\left[\frac{1}{2}\bar{b}_m\partial_m a_{ll'}+\frac{1}{2}\sigma_{l'i}a_{mk}\partial_{mk}\sigma_{li}-a_{kl'}\partial_k\bar{b}_l\right]+\left[\sigma_{l'i}a_{mk}-\sigma_{ki}a_{ml'}\right]\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho\Big\}.\end{aligned}$$

Remark 3.4 *The form of the term* $tr(\Lambda_\delta\Gamma)$ *in the above proposition is inspired from the term* $tr(\mathbf{XY})$ *in [2] pp 163-164 where* $\mathbf{X}=2\Lambda_\delta$*. One has*

$$\begin{aligned}\Gamma_{12}=\;(\nabla^*\rho\;a)_j\;\partial_j(\sigma_{ki}\partial_k\rho)\sigma_{li}\partial_l\rho=&\frac{1}{2}(\nabla^*\rho\;a)_j\left[\partial_j(\sigma_{ki}\partial_k\rho)\sigma_{li}\partial_l\rho+\partial_j(\sigma_{li}\partial_l\rho)\sigma_{ki}\partial_k\rho\right]\\=&\frac{1}{2}(\nabla^*\rho\;a)_j\partial_j\left[\partial_l\rho a_{kl}\partial_k\rho\right]=\frac{1}{2}(\nabla^*\rho\;a)\nabla(\nabla^*\rho a\nabla\rho)\end{aligned}$$

which, with $\frac{\partial F}{\partial x}:=(\frac{\partial F_i}{\partial x_j})_{i,j}$ *denoting the Jacobian matrix, equals*

$$\begin{aligned}\frac{1}{2}(\nabla^*\rho\;a)_j\partial_j\left[\partial_l\rho a_{kl}\partial_k\rho\right]=&\frac{1}{2}(\nabla^*\rho\;a)_j\;(\partial_{kj}\rho\;a_{kl}\;\partial_l\rho+\partial_j\left[a_{kl}\;\partial_l\rho\right]\partial_k\rho)\\=&\frac{1}{2}\nabla^*\rho\;a\frac{\partial(\nabla\rho)}{\partial x}a\nabla\rho+\frac{1}{2}\nabla^*\rho\;a\frac{\partial(a\nabla\rho)}{\partial x}^*\nabla\rho\end{aligned}$$

and corresponds to $4\mathbf{Y}_{12}$ in [2], p. 164. Similarly, $\Gamma_{22} = 4\mathbf{Y}_{22}$. However Γ_{11} cannot in general be identified with $4\mathbf{Y}_{11}$. For instance, in the case of scalar diffusion $\mathbf{D}(x) = a(x)/2 = D(x)I_d$ for some real valued function D, the term $\Gamma_{11}(x)$ above when written in terms of D reads

$$|\nabla D|^2|\nabla\rho|^2 + 4D\partial_j D\partial_i\rho\partial_{ij}\rho + 4D^2\sum_{ij}(\partial_{ij}\rho)^2,$$

for the choice $\sigma(x) = \sqrt{D(x)}I_d$, whereas

$$4\mathbf{Y}_{11} = 4\left(D^2\sum_{ij}(\partial_{ij}\rho)^2 + \left(\frac{n}{4}-\frac{1}{2}\right)(\nabla\rho\cdot\nabla D)^2 + 2D\partial_j D\partial_i\rho\partial_{ij}\rho - D(\nabla\rho\cdot\nabla D)\triangle\rho + \frac{1}{2}|\nabla D|^2|\nabla\rho|^2\right).$$

Moreover, our term Γ_{11} above is non-intrinsic, in the sense that it cannot in general be written in terms of the diffusion matrix a only (without making explicit use of σ), contrary to the term $\mathbf{Y}_{11}$ in the matrix of [2].

Before proving Proposition 3.3, following [1], we introduce an additional assumption on U that will be made in all the sequel :

$H7'$) $\forall r \in (0,\infty)$, $(U^{(3)}(r))^2 \leq \frac{1}{2}U''(r)U^{(4)}(r)$.

By $H7'$), Λ_δ is a positive semidefinite matrix. Since by Cauchy Schwarz inequality,

$$\begin{aligned}((\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\ a\nabla(\sigma_{\bullet i}\cdot\nabla\rho))^2 &= ((\sigma_{\bullet i}\cdot\nabla\rho)\sigma^*\nabla\rho.\sigma^*\nabla(\sigma_{\bullet i}\cdot\nabla\rho))^2\\ &\leq \sum_i(\sigma_{\bullet i}\cdot\nabla\rho)^2|\sigma^*\nabla\rho|^2\sum_i|\sigma^*\nabla(\sigma_{\bullet i}\cdot\nabla\rho)|^2\\ &= |\nabla^*\rho a\nabla\rho|^2\times\nabla^*(\sigma_{\bullet i}\cdot\nabla\rho)a\nabla(\sigma_{\bullet i}\cdot\nabla\rho).\end{aligned}$$

the determinant of the matrix Γ is nonnegative, and this matrix also is positive semidefinite. As an easy consequence we have

Corollary 3.5

$$\forall\delta>0,\ d\left[U_\delta''(\rho)\nabla^*\rho a\nabla\rho\right] \geq U_\delta''(\rho)\bar{\theta}dt + d\hat{M}^{(\delta)}.$$

Notice that one could preserve the positive semidefiniteness of the matrix Γ when replacing Γ_{11} by the smaller coefficient $\sum_{i=1}^{d'}\left(\nabla^*\rho\ a\nabla(\sigma_{\bullet i}\cdot\nabla\rho)\right)^2/|\nabla^*\rho a\nabla\rho|$, which amounts to replace the squared norms of the vectors $\sigma^*\nabla(\sigma_{\bullet i}\cdot\nabla\rho)$ by the ones of their orthogonal projection on $\sigma^*\nabla\rho$. Unfortunately, we have not been able to take advantage of this possibility.

Proof of Proposition 3.3.

By Itô 's formula we get $d\sigma_{li} = [\sigma_{mr}\partial_m\sigma_{li}]\,d\bar{W}_t^r + \left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right]dt$. We then have

$$\begin{aligned}d\left[\sigma_{li}\partial_l\rho\right] &= \sigma_{li}d\partial_l\rho + \partial_l\rho d\sigma_{li} + d\langle\partial_l\rho,\sigma_{li}\rangle\\ &= \partial_k\left[\partial_l\rho\sigma_{li}\right]\sigma_{kr}d\bar{W}^r + \partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right] - \sigma_{li}\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr} + \partial_k\rho\partial_l\bar{b}_k\right]\\ &\quad+\ a_{mk}\partial_{lk}\rho\partial_m\sigma_{li}\end{aligned}$$

where we used in the stochastic integral the fact that $\partial_l\rho\sigma_{mr}\partial_m\sigma_{li}+\sigma_{li}\sigma_{kr}\partial_{lk}\rho=\partial_l\rho\sigma_{kr}\partial_k\sigma_{li}+\sigma_{li}\sigma_{kr}\partial_{lk}\rho=\partial_k\left[\partial_l\rho\sigma_{li}\right]\sigma_{kr}$. It follows that

$$\begin{aligned}
d\left[\nabla^*\rho a\nabla\rho\right] &= d\left[\sigma_{li}\partial_l\rho\ \sigma_{l'i}\partial_{l'}\rho\right]\\
&= 2\ \sigma_{l'i}\ \partial_{l'}\rho\partial_k\left[\sigma_{li}\partial_l\rho\right]\sigma_{kr}d\bar{W}^r +\ 2\Big\{\ \left[\sigma_{l'i}\partial_{l'}\rho a_{mk}\partial_m\sigma_{li}\partial_{lk}\rho\right]\\
&\quad+\sigma_{l'i}\partial_{l'}\rho\partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li}+\frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right]-a_{ll'}\partial_{l'}\rho\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr}+\partial_k\rho\partial_l\bar{b}_k\right]\Big\}dt\\
&\quad+a_{kk'}\partial_k\left[\partial_l\rho\sigma_{li}\right]\partial_{k'}\left[\partial_{l'}\rho\sigma_{l'i}\right]dt
\end{aligned}$$

On the other hand, using (3.3) at $x=Y_0$ we have $dU''_\delta(\rho)=U^{(3)}_\delta(\rho)\sigma_{nr}\partial_n\rho\ d\bar{W}^r+\frac{1}{2}U^{(4)}_\delta(\rho)a_{nj}\partial_n\rho\partial_j\rho\ dt$ which combined with the previous expression yields

$$\begin{aligned}
d\left[U''_\delta(\rho)\nabla^*\rho a\nabla\rho\right] =&2U''_\delta(\rho)\Big\{\ \left[\sigma_{l'i}\partial_{l'}\rho a_{mk}\partial_m\sigma_{li}\partial_{lk}\rho\right]+\sigma_{l'i}\partial_{l'}\rho\partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li}+\frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right]\\
&-a_{ll'}\partial_{l'}\rho\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr}+\partial_k\rho\partial_l\bar{b}_k\right]\Big\}dt\ +\ d\hat{M}^{(\delta)}\\
&+\ U''_\delta(\rho)a_{kk'}\partial_k\left[\partial_l\rho\sigma_{li}\right]\partial_{k'}\left[\partial_{l'}\rho\sigma_{l'i}\right]dt+\frac{1}{2}U^{(4)}_\delta(\rho)\left|\nabla^*\rho a\nabla\rho\right|^2dt\\
&+\ 2U^{(3)}_\delta(\rho)\sigma_{l'i}\partial_{l'}\rho\partial_k\left[\sigma_{li}\partial_l\rho\right]a_{jk}\partial_j\rho dt.
\end{aligned}$$

Equivalently,

$$\begin{aligned}
d\left[U''_\delta(\rho)\nabla^*\rho a\nabla\rho\right] =&2U''_\delta(\rho)\Big\{\ \left[\sigma_{l'i}\partial_{l'}\rho a_{mk}\partial_m\sigma_{li}\partial_{lk}\rho\right]+\sigma_{l'i}\partial_{l'}\rho\partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li}+\frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right]\\
&-a_{ll'}\partial_{l'}\rho\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr}+\partial_k\rho\partial_l\bar{b}_k\right]\Big\}dt\ +\ d\hat{M}^{(\delta)}\\
&+\Big(\ U''_\delta(\rho)\nabla^*(\sigma_{\bullet i}\cdot\nabla\rho)a\nabla(\sigma_{\bullet i}\cdot\nabla\rho)+\frac{1}{2}U^{(4)}_\delta(\rho)\left|\nabla^*\rho a\nabla\rho\right|^2\\
&+\ 2U^{(3)}_\delta(\rho)(\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\ a\nabla(\sigma_{\bullet i}\cdot\nabla\rho)\Big)dt
\end{aligned}\tag{3.8}$$

∎

We recall properties of the function U pointed out in [1] (see Remark 2.3 therein) which will be used in proving the following results.

Remark 3.6 *Condition $H7'$) implies that $\left(\frac{1}{U''}\right)''\leq 0$ at points where $U''\neq 0$. Since $U''\geq 0$, and excluding the uninteresting case where U'' identically vanishes, the previous implies that $\frac{1}{U''}$ is finite in $[0,\infty)$, and therefore that U is strictly convex. We then deduce from $H7'$) that $U^{(4)}\geq 0$ in $(0,\infty)$. By concavity and positivity of $\frac{1}{U''}$ this function is moreover non decreasing, and we deduce that $U^{(3)}\leq 0$ in $(0,\infty)$.*

We introduce one last assumption on the density flow ρ_t:

$H6')_{p_0}$ For each $T'\in(0,T]$ the following integrals are finite:

- $\int_0^{T'}\left|U^{(3)}(\rho)\vee -1\right|^2\left|\nabla^*\rho a\nabla\rho\right|^3 p_\infty(x)dxdt$

- $\int_0^{T'} \left(U''(\rho) \wedge 1\right)^2 \nabla^*(\nabla^*\rho a\nabla\rho)a\nabla(\nabla^*\rho a\nabla\rho)p_\infty(x)dxdt$
- $\int_0^{T'}(U''(\rho)\wedge 1)\big[\,|(\sigma_{l'i}a_{m\bullet} - \sigma_{\bullet i}a_{ml'})\partial_m\sigma_{li}|+|\partial_k\left([\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}]\,\partial_m\sigma_{li}\right)|\,\big]|\partial_{l'}\rho||\partial_l\rho|p_\infty(x)dxdt$
- $\int_0^{T'}(U''(\rho) \wedge 1)\big[\,|(\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'})\partial_m\sigma_{li}(\partial_l\rho\partial_k \ln p_\infty + \partial_{lk}\rho)|\,\big]|\partial_{l'}\rho|p_\infty(x)dxdt$

Notice that in the case that $a \geq cI_d$ for some $c > 0$, the third integral converges always as it can be upper bounded by the Fisher information.

Theorem 3.7 *Let Θ denote the $d \times d$ matrix defined by*

$$\begin{aligned}\Theta_{ll'} = \sigma_{l'i}[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}] - a_{kl'}\partial_k\bar{b}_l + (\sigma_{ki}a_{ml'} - \sigma_{l'i}a_{mk})\partial_m\sigma_{li}\partial_k \ln(p_\infty)\\ + \partial_k[(\sigma_{ki}a_{ml'} - \sigma_{l'i}a_{mk})\partial_m\sigma_{li}]\end{aligned}$$

and assume that the symmetric matrix $(\Theta + \Theta^*)(t, x)$ is $p_\infty(x)dxdt - a.e.$ positive semidefinite. Then, for a.e. $t \in [0, T]$ one has

$$\frac{d}{dt}\int_{\rho_t>0} U''(\rho_t)[\nabla^*\rho_t a\nabla\rho_t]p_\infty dx \geq \int_{\rho_t>0} U''(\rho_t)\nabla^*\rho_t(\Theta + \Theta^*)\nabla\rho_t p_\infty dx. \tag{3.9}$$

If moreover, $H_U(p_s|p_\infty)$ is finite for some $s \geq 0$ and the diffusion matrix a is locally uniformly strictly positive definite, then $H_U(p_t|p_\infty)$ converges to 0 as $t \to \infty$.

Proof . Let us first suppose that (3.9) holds and deduce the last assertion. Reverting time in (3.9), we obtain that $t \mapsto I_U(p_t|p_\infty)$ is non-increasing. When $H_U(p_s|p_\infty)$ is finite for some $s \geq 0$, writing (2.7) on the interval $[s, T]$ in place of $[0, T]$, we deduce that $I_U(p_t|p_\infty)$ tends to 0 as $t \to \infty$. When a is locally uniformly strictly positive definite, the beginning of the proof of Theorem 2.5 (before Part(a)) [2], ensures that p_t tends to p_∞ in $L^1(\mathbb{R}^d)$. As a consequence, in the notations of Proposition 1.1, $\mathbb{E}\left|\frac{dP_t}{dQ_t}(X_t^{Q_0}) - 1\right|$ tends to 0 as $t \to \infty$ and therefore the a.s. limit of $\frac{dP_t}{dQ_t}(X_t^{Q_0})$ is equal to 1. By corollary 1.2, one concludes that $H_U(p_t|p_\infty)$ tends to 0.

Let us now check (3.9). Since U'' is continuous and non increasing in $(0, \infty)$ by Remark 3.6, one has $U''_\delta(r) \nearrow U''(r)$ for each $r > 0$ as $\delta \to 0$. It is therefore enough to obtain (the integrated version of) inequality (3.9) with U''_δ instead of U'', as monotone convergence allows us to pass to the limit as $\delta \to 0$ on both sides. For $0 \leq r \leq t < T$ we have by Corollary 3.5 that

$$\begin{aligned}&[U''_\delta(\rho)\nabla^*\rho a\nabla\rho](t, Y_t) - [U''_\delta(\rho)\nabla^*\rho a\nabla\rho](r, Y_r)\\ &\qquad \geq \hat{M}^{(\delta)}_t - \hat{M}^{(\delta)}_r + 2\int_r^t U''_\delta(\rho)\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}]\,\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho ds\\ &\qquad + 2\int_r^t U''_\delta(\rho)\partial_{l'}\rho\partial_l\rho\left(\sigma_{l'i}\left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right] - a_{ml'}\partial_m\bar{b}_l\right)ds.\end{aligned}$$

Since $\partial_{kl'}\rho U''_\delta(\rho)\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}] = 0$ and

$$\partial_k(U''_\delta(\rho))\partial_{l'}\rho\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}] = U^{(3)}_\delta(\rho)\partial_k\rho\partial_{l'}\rho\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}] = 0,$$

one has

$$\begin{aligned}U''_\delta(\rho)\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}]\,\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho =& \frac{1}{p_\infty}\partial_k\left(\partial_l\rho\partial_{l'}\rho U''_\delta(\rho)\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}]\,\partial_m\sigma_{li}\,p_\infty\right)\\ &- \partial_l\rho\partial_{l'}\rho U''_\delta(\rho)\,[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}]\,\partial_m\sigma_{li}\partial_k \ln p_\infty\\ &- \partial_l\rho\partial_{l'}\rho U''_\delta(\rho)\partial_k\left([\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}]\,\partial_m\sigma_{li}\right).\end{aligned} \tag{3.10}$$

We deduce that

$$
\begin{aligned}
&[U_\delta''(\rho)\nabla^*\rho a\nabla\rho](t,Y_t)-[U_\delta''(\rho)\nabla^*\rho a\nabla\rho](r,Y_r)\\
&\geq \hat{M}_t^{(\delta)}-\hat{M}_r^{(\delta)}+2\int_r^t U_\delta''(\rho)\Theta_{ll'}\partial_{l'}\rho\partial_l\rho\,ds+2\int_r^t\frac{1}{p_\infty}\partial_k\left(\partial_l\rho\partial_{l'}\rho U_\delta''(\rho)\left[\sigma_{l'i}a_{mk}-\sigma_{ki}a_{ml'}\right]\partial_m\sigma_{li}\,p_\infty\right)ds.
\end{aligned}
\tag{3.11}
$$

Now, the quadratic variation of $\hat{M}^{(\delta)}$ is bounded above in $[0,T)$ by a constant times

$$
\int_0^t\left[|U_\delta^{(3)}(\rho)|^2|\nabla^*\rho a\nabla\rho|^3(Y_s)+\left(U_\delta''(\rho)\right)^2\nabla^*(\nabla^*\rho a\nabla\rho)a\nabla(\nabla^*\rho a\nabla\rho)\right](Y_s)ds.
$$

This and our assumptions imply that $\hat{M}^\delta$ is a martingale in $[0,T)$ for all $\delta>0$ sufficiently small. Indeed, we have from Remark 3.6 that $U_\delta''(r)\leq U''(\delta)\wedge U''(r)$ and $|U_\delta^{(3)}(r)|\leq|U^{(3)}(\delta)|\wedge|U^{(3)}(r)|$ for all $r\geq 0$. Therefore (since $U''>0$) we have $U_\delta''(r)\leq(U''(r)\wedge 1)\mathbf{1}_{U''(\delta)\leq 1}+U''(\delta)(U''(r)/U''(\delta))\wedge 1)\mathbf{1}_{U''(\delta)>1}$ whence $U_\delta''(r)\leq(U''(\delta)+1)(U''(r)\wedge 1)$. As $U^{(3)}$ is non decreasing and non positive, either $|U^{(3)}(\delta)|\neq 0$ for all δ sufficiently small, in which case we similarly get $|U_\delta^{(3)}(r)|\leq(|U^{(3)}(\delta)|+1)(|U^{(3)}(r)|\wedge 1)$, or otherwise $U_\delta^{(3)}$ identically vanishes for all δ. Assumption $H6')_{p_\infty}$ and the previous then ensure that $\langle M^{(\delta)}\rangle_t$ has finite expectation for $t\in[0,T)$.

In order to conclude that inequality (3.9) holds for the function U_δ , noting that $\nabla\rho_t$ vanishes on $\{\rho_t=0\}$, it is enough to show that the last integral in (3.11) has (well defined) null expectation. Using (3.10) and Assumption $H6')_{p_\infty}$ we obtain (with the same control for $U_\delta''(r)$ as before) that

$$
\begin{aligned}
&\mathbb{E}_\infty^{T\to 0}\int_r^t\left|\frac{1}{p_\infty}\partial_k\left(\partial_l\rho\partial_{l'}\rho U_\delta''(\rho)\left[\sigma_{l'i}a_{mk}-\sigma_{ki}a_{ml'}\right]\partial_m\sigma_{li}\,p_\infty\right)\right|(Y_s)ds\\
&\qquad=\int_r^t\int_{\mathbb{R}^d}\left|\partial_k\left(\partial_l\rho\partial_{l'}\rho U_\delta''(\rho)\left[\sigma_{l'i}a_{mk}-\sigma_{ki}a_{ml'}\right]\partial_m\sigma_{li}\,p_\infty\right)\right|dxds<\infty
\end{aligned}
\tag{3.12}
$$

which shows that the expectation of the last term in (3.11) is well defined. Moreover, the (everywhere defined) spatial divergence of $g(s,x):=\partial_l\rho_s\partial_{l'}\rho_s U_\delta''(\rho_s)\left[\sigma_{l'i}a_{m\bullet}-\sigma_{\bullet i}a_{ml'}\right]\partial_m\sigma_{li}\,p_\infty$ is $L^1(dx,\mathbb{R}^d)$ for a.e. s. For such s and $\phi_n\in C_0^\infty(\mathbb{R}^d)$ such that $0\leq\phi_n\leq 1$, $0\leq|\nabla\phi_n|\leq 1$, $\phi_n(x)=1$ for $x\in B(0,n)$ and $\phi_n(x)=0$ for $x\in B(0,2n)^c$,

$$
0=\int_{\mathbb{R}^d}\nabla.(\phi_n(x)g(s,x))dx=\int_{\mathbb{R}^d}\phi_n(x)\nabla.g(s,x)dx+\int_{\mathbb{R}^d}\nabla\phi_n(x).g(s,x)dx.
$$

Since by Lebesgue's theorem, the second term of the right-hand-side tends to 0 as $n\to\infty$, the limit $\int_{\mathbb{R}^d}\nabla.g(s,x)dx$ of the first term is equal to 0. ∎

Theorem 3.8 *Under the hypotheses of Theorem 3.7 and if the matrix* Θ *satisfies the non-intrinsic Bakry-Emery criterion*

NIBEC) $\exists\lambda>0,\ \forall x\in\mathbb{R}^d,\ \frac{1}{2}(\Theta+\Theta^*)(x)\geq\lambda a(x).$

then the non-increasing function $t\mapsto H_U(p_t|p_\infty)$ *converges at exponential rate* 2λ *to its limit as* $t\to\infty$*. When, moreover, the diffusion matrix* a *is locally uniformly strictly positive definite,*

then this limit is equal to 0 *as soon as* $H_U(p_s|p_\infty)$ *is finite for some* $s \geq 0$ *and the convex Sobolev inequality*

$$H_U(p|p_\infty) \leq \frac{1}{2\lambda} I_U(p|p_\infty) \tag{3.13}$$

holds for any probability density p *on* $\mathbb{R}^d$.

Remark 3.9 *In view of Remark 2.15, the local uniform strict positive definiteness assumption on the diffusion matrix* a *may be replaced by some hypoellipticity assumption for the convex Sobolev inequality* (3.13) *to hold for any probability density* p *on* $\mathbb{R}^d$.

Proof . Reverting time in (3.9) and using *NIBEC)*, one obtains

$$\frac{d}{ds} I_U(p_s|p_\infty) \leq -2\lambda I_U(p_s|p_\infty).$$

Hence $\forall s \geq 0,\ I_U(p_s|p_\infty) \leq e^{-2\lambda s} I_U(p_0|p_\infty)$. Since by Theorem 2.9, one has $\frac{d}{ds} H_U(p_s|p_\infty) = -I_U(p_s|p_\infty)$, one deduces that

$$0 \leq H_U(p_s|p_\infty) - \lim_{t\to\infty} H_U(p_t|p_\infty) = \int_s^\infty I_U(p_s|p_\infty) \leq \frac{e^{-2\lambda s}}{2\lambda} I_U(p_0|p_\infty).$$

When a is locally uniformly strictly positive definite, if $H_U(p_0|p_\infty) < +\infty$, then $\lim_{t\to\infty} H_U(p_t|p_\infty) = 0$ by Theorem 3.7. Moreover, the convex Sobolev inequality for $p = p_0$ is just the last inequality written for $s = 0$. It can be extended to arbitrary probability densities p on $\mathbb{R}^d$ by simple closure.

∎

Remark 3.10 *i) Notice that* $\frac{1}{2}(\Theta + \Theta^*)$ *rewrites as*

$$\begin{aligned}
&\frac{1}{2}\bar{b}_m \partial_m a_{ll'} - \frac{1}{2}(a_{kl'}\partial_k \bar{b}_l + a_{kl}\partial_k \bar{b}_{l'}) + \frac{1}{4} a_{mk}\partial_{mk} a_{ll'} - \frac{1}{2} a_{mk}\partial_m \sigma_{li}\partial_k \sigma_{l'i} \\
&+ \frac{1}{2}\sigma_{ki}(\partial_m \sigma_{li} a_{ml'} + \partial_m \sigma_{l'i} a_{ml})\partial_k \ln(p_\infty) - \frac{1}{2} a_{mk}\partial_m a_{ll'}\partial_k \ln(p_\infty) \\
&+ \frac{1}{2}\partial_k[\sigma_{ki}(\partial_m \sigma_{li} a_{ml'} + \partial_m \sigma_{l'i} a_{ml}) - a_{mk}\partial_m a_{ll'}] \\
= &- \frac{1}{2} b_m \partial_m a_{ll'} + \frac{1}{2}(a_{kl'}\partial_k b_l + a_{kl}\partial_k b_{l'}) - \frac{1}{4} a_{mk}\partial_{mk} a_{ll'} - \frac{1}{2}(a_{kl'}\partial_{kj} a_{lj} + a_{kl}\partial_{kj} a_{l'j}) \\
&- a_{kl} a_{jl'}\partial_{kj} \ln(p_\infty) - \frac{1}{2}(a_{kl}\partial_k a_{l'j} + a_{kl'}\partial_k a_{lj})\partial_j \ln(p_\infty) - \frac{1}{2} a_{mk}\partial_m \sigma_{li}\partial_k \sigma_{l'i} \\
&+ \frac{1}{2}\sigma_{ki}(\partial_m \sigma_{li} a_{ml'} + \partial_m \sigma_{l'i} a_{ml})\partial_k \ln(p_\infty) + \frac{1}{2}\partial_k[\sigma_{ki}(\partial_m \sigma_{li} a_{ml'} + \partial_m \sigma_{l'i} a_{ml})],
\end{aligned} \tag{3.14}$$

both being non-intrisic expressions that cannot be rewritten without making use of the square root σ. *Since we have got rid of the nonnegative term* $tr(\Lambda_\delta \Gamma)$ *which appears in the first equation in Proposition 3.3 and involves the non-intrisic term* Γ_{11}, *it is natural that we obtain a non-intrisic Bakry Emery criterion.*

ii) In case $\sigma = \sqrt{2\nu} I_d$ *and* $b = -(\nabla V + F)$ *with* F *such that* $\nabla.(e^{-V/\nu}F) = 0$, *then* $p_\infty \propto e^{-V/\nu}$, $\bar{b} = -b + 2\nu\nabla \ln p_\infty = -\nabla V + F$ *and* $\Theta = 2\nu(\nabla^2 V - \nabla F)$. *Therefore condition NIBEC) writes* $\exists \lambda > 0,\ \forall x \in \mathbb{R}^d,\ \nabla^2 V(x) - \frac{\nabla F + \nabla F^*}{2}(x) \geq \lambda I_d$ *which is exactly condition (A2) in the introduction of [2], page 158.*

4 An example

The next example shows that our criterion and an appropriate choice of the square root of the diffusion matrix can ensure the exponential convergence to equilibrium when the classic Bakry Emery criterion fails.

We consider a reversible diffusion process in $\mathbb{R}^d$ with $d = 2$, such that for each $(x_1, x_2) \in \mathbb{R}^2$,

$$a(x_1, x_2) = I_2, \quad \text{and} \quad b(x_1, x_2) = -\nabla V(x_1, x_2)$$

where, for some $\alpha \in (0, 1)$, V is the convex potential

$$V(x_1, x_2) := |x_1|^2 + v(|x_1 - x_2|) + v(|x_2|) \text{ with } v(z) = \begin{cases} z^{2+\alpha} \text{ if } z \in [0, 1] \\ 1 + (2+\alpha)(z-1)\frac{(1+\alpha)z + (1-\alpha)}{2} \text{ if } z \geq 1 \end{cases}.$$

The choice of v ensures that V is globally C^2 and quadratic far from the origin, which implies that b is globally Lipschitz continuous. The invariant measure is in this case $p_\infty \propto e^{-2V}$, and we have

$$\begin{aligned}
\partial_1 V =& 2x_1 + (2+\alpha) sign(x_1 - x_2)[|x_1 - x_2|^{1+\alpha} \mathbf{1}_{\{|x_1 - x_2| \leq 1\}} + ((1+\alpha)|x_1 - x_2| - \alpha)\mathbf{1}_{\{|x_1 - x_2| > 1\}}] \\
\partial_2 V =& (2+\alpha) sign(x_2)[|x_2|^{1+\alpha} \mathbf{1}_{\{|x_2| \leq 1\}} + ((1+\alpha)|x_2| - \alpha)\mathbf{1}_{\{|x_2| > 1\}}] \\
&+ (2+\alpha) sign(x_2 - x_1)[|x_2 - x_1|^{1+\alpha} \mathbf{1}_{\{|x_2 - x_1| \leq 1\}} + ((1+\alpha)|x_2 - x_1| - \alpha)\mathbf{1}_{\{|x_2 - x_1| > 1\}}]
\end{aligned}$$

and

$$\nabla^2 V = \begin{pmatrix} 2 & 0 \\ 0 & (2+\alpha)(1+\alpha)(|x_2| \wedge 1)^\alpha \end{pmatrix} + (2+\alpha)(1+\alpha)(|x_1 - x_2| \wedge 1)^\alpha \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Notice that the classic Bakry-Emery criterion fails in this case since $\nabla^2 V(0, 0)$ is singular. We consider moreover a square root σ of the identity matrix of the form

$$\sigma = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

for a function $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ of class C^2 to be chosen later. Starting from (3.14), we obtain after some computations

$$\frac{1}{2}(\Theta + \Theta^*) = \nabla^2 V - \frac{1}{2}|\nabla\phi|^2 I_2 + \begin{pmatrix} \partial_{12}\phi & \frac{\partial_{22}\phi - \partial_{11}\phi}{2} \\ \frac{\partial_{22}\phi - \partial_{11}\phi}{2} & -\partial_{12}\phi \end{pmatrix} + \begin{pmatrix} -2\partial_1\phi\partial_2 V & \partial_1\phi\partial_1 V - \partial_2\phi\partial_2 V \\ \partial_1\phi\partial_1 V - \partial_2\phi\partial_2 V & 2\partial_2\phi\partial_1 V \end{pmatrix}$$

We now consider a parameter $\varepsilon > 0$ which will be chosen small and a C^2 function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\varphi(s) = s$ if $|s| \leq 1$ and $\varphi(s) = 0$ if $|s| \geq 2$. Then, we define

$$\phi(x_1, x_2) = -\varepsilon\varphi_\varepsilon(x_1)\varphi_\varepsilon(x_2), \quad (x_1, x_2) \in \mathbb{R}^2$$

where $\varphi_\varepsilon(s) = \varepsilon\varphi(s/\varepsilon)$. Notice that

$$\varphi_\varepsilon = O(\varepsilon), \quad \varphi_\varepsilon'' = O(1/\varepsilon), \quad \text{and } \varphi_\varepsilon' = \begin{cases} 1 & \text{if } |s| \leq \varepsilon, \\ O(1) & \text{if } \varepsilon < |s| < 2\varepsilon, \\ 0 & \text{if } |s| \geq 2\varepsilon. \end{cases}$$

Then, defining $B_\varepsilon := \{(x_1, x_2) \in \mathbb{R}^2 \; s.t. \, |x_1| \vee |x_2| \leq \varepsilon\}$ and $C_\varepsilon := B_{2\varepsilon} \backslash B_\varepsilon$, we have

$$\partial_1\phi(x_1, x_2), \partial_2\phi(x_1, x_2) = \begin{cases} O(\varepsilon^2) & \text{if } (x_1, x_2) \in B_{2\varepsilon}, \\ 0 & \text{if } (x_1, x_2) \in B_{2\varepsilon}^c, \end{cases}$$

$$\partial_{12}\phi(x_1,x_2)=\begin{cases}-\varepsilon & \text{if } (x_1,x_2)\in B_\varepsilon,\\ O(\varepsilon) & \text{if } (x_1,x_2)\in C_\varepsilon,\\ 0 & \text{if } (x_1,x_2)\in B^c_{2\varepsilon},\end{cases}$$

$$\frac{1}{2}(\partial_{11}\phi(x_1,x_2)-\partial_{22}\phi(x_1,x_2))=\begin{cases}0 & \text{if } (x_1,x_2)\in B_\varepsilon,\\ O(\varepsilon) & \text{if } (x_1,x_2)\in C_\varepsilon,\\ 0 & \text{if } (x_1,x_2)\in B^c_{2\varepsilon},\end{cases}$$

and $\partial_1 V=O(\varepsilon),\partial_2 V=O(\varepsilon^{1+\alpha})$ on $B_{2\varepsilon}$. It follows that

$$\frac{1}{2}(\Theta+\Theta^*)=\nabla^2 V+\begin{pmatrix}-\varepsilon & 0\\ 0 & \varepsilon\end{pmatrix}+O(\varepsilon^3)\geq\begin{pmatrix}2-\varepsilon & 0\\ 0 & \varepsilon\end{pmatrix}+O(\varepsilon^3)\quad\text{ on } B_\varepsilon.$$

Next, the smallest eigenvalue of $\nabla^2 V(x_1,x_2)$, is given by

$$\gamma_-:=1+\kappa_1+\kappa_2/2-\sqrt{1+\kappa_1^2-\kappa_2+\kappa_2^2/4}\geq 0$$

with $\kappa_1=\kappa_1(x_1,x_2):=(2+\alpha)(1+\alpha)(|x_1-x_2|\wedge 1)^\alpha$ and $\kappa_2=\kappa_2(x_1,x_2):=(2+\alpha)(1+\alpha)(|x_2|\wedge 1)^\alpha$. Since $\gamma_-=\kappa_1+\kappa_2+O(\kappa_1^2+\kappa_2^2)$ as $\kappa_1^2+\kappa_2^2\to 0$ and $|x_2|^\alpha+|x_1-x_2|^\alpha\geq(|x_2|+|x_1-x_2|)^\alpha\geq|x_1|^\alpha$, one deduces that on C_ε,

$$\frac{1}{2}(\Theta+\Theta^*)=\nabla^2 V+O(\varepsilon)\geq(2+\alpha)(1+\alpha)\varepsilon^\alpha I_2+O(\varepsilon).$$

Last, since κ_1 and κ_2 are continuous and bounded functions of (x_1,x_2), and γ_- is a continuous function of (κ_1,κ_2) only vanishing at the origin, $\inf_{(x_1,x_2)\in B^c_{2\varepsilon}}\gamma_->0$. One concludes that for ε small enough *NIBEC)* holds.

Remark 4.1

- *The potential V is a particular case of the examples considered by Arnold, Carlen and Ju in the Section 3 of [2]. But they first modify the Fokker-Planck equation by adding a non-symmetric drift term F like in Remark 3.10 ii) to check that p_∞ satisfies the convex Sobolev inequality* (3.13). *Exponential convergence to* 0 *of $H_U(p_t|p_\infty)$ for the solution p_t of the original Fokker-Planck equation is only deduced in a second step. With our non-intrisic Bakry Emery criterion, we are able to prove without considering a modified partial differential equation that p_∞ satisfies the convex Sobolev inequality* (3.13) *and that $H_U(p_t|p_\infty)$ converges exponentially to* 0. *We modify the stochastic differential equation but not the law of its solution.*

- *We have supposed that V is quadratic far from the origin to ensure that b satisfies H4). But the boundedness assumption on the first order derivatives of b in H4) may be relaxed to local boundedness when conditions ensuring existence for the SDE and preservation of the diffusion property under time reversal are added. In the case of constant diffusion $a(x)=I_d$ with drift $b(x)=-\nabla V(x)$ for a nonnegative C^2 potential V, the following hypotheses on the behaviour of V at infinity:*

$$\limsup_{|x|\to\infty}\frac{-x^*\nabla V(x)}{|x|^2}<+\infty,\ \limsup_{|x|\to\infty}\frac{\Delta V}{|\nabla V|^2}(x)<2\ \textit{and}\ \limsup_{|x|\to\infty}\frac{\sqrt{\partial_{ik}V\partial_{ik}V}}{V}(x)=0 \quad (4.1)$$

provide such sufficient additional conditions for the SDE $dX_t=\sigma(X_t)dW_t-\nabla V(X_t)dt$, when $\sigma(x)$ is any globally Lipschitz continuous choice of the square root of the identity. Indeed, by computing $d|X_t|^2$, one sees that the first condition prevents explosion for the SDE which has locally Lipschitz coefficients. Since for $c>0$,

$$de^{cV(X_t)}=e^{cV(X_t)}\left(\nabla^* V(X_t)\sigma(X_t)dW_t+\frac{c}{2}[\Delta V+(c-2)|\nabla V|^2](X_t)dt\right),$$

the second condition ensures that for c small enough, $t \mapsto \mathbb{E}(e^{cV(X_t)})$ is locally bounded when $\mathbb{E}(e^{cV(X_0)}) < +\infty$. With the inequality

$$\mathbb{E}\left(\exp(4\int_0^T \sqrt{\partial_{ik}V\partial_{ik}V(X_t)}dt)\right) \leq K\mathbb{E}\left(\exp(\frac{c}{T}\int_0^T V(X_t)dt)\right)$$

deduced from the third assumption, one concludes by Jensen's inequality that $\mathbb{E}\left(\exp(4\int_0^T \sqrt{\partial_{ik}V\partial_{ik}V(X_t)}dt)\right)$ is finite as soon as $\mathbb{E}(e^{cV(X_0)}) < +\infty$. Hence condition (3.9) in Theorem 3.3 of [11], which is enough for Theorem 2.3 to hold, is satisfied. Notice that (4.1) *is satisfied for instance by the potential $V(x_1, x_2) = x_1^2 + |x_1 - x_2|^{2+\alpha} + |x_2|^{2+\alpha}$ from which the potential in the above example was derived by replacing the super-quadratic terms by quadratic ones outside the unit ball.*

References

[1] Arnold, A., Markowich, P., Toscani, G., Unterreiter, A., : *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations*, Comm. Partial Differential Equations 26 (2001), no. 1-2, 43-100.

[2] Arnold, A., Carlen, E. and Ju, Q. : *Large-time behavior of non-symmetric Fokker-Planck type equations*, Communications on Stochastic Analysis, 2(1), 153–175, 2008.

[3] Cattiaux, P. :*A pathwise approach of some classical inequalities.* Potential Anal. 20 (2004), no. 4, 361–394

[4] Föllmer, H.: *Time reversal on Wiener space* Stochastic processes—mathematics and physics (Bielefeld, 1984), 119–129, Lecture Notes in Math., 1158, Springer, Berlin, 1986.

[5] Friedman, A .: *Partial differential equations of parabolic type.* Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964

[6] Haussmann, U. G.; Pardoux, É.: *Time reversal of diffusions* Ann. Probab. 14 (1986), no. 4, 1188–1205.

[7] Hwang, C.-R. ; Hwang-Ma, S.-Y. ; Sheu, S.-J. Accelerating diffusions. Ann. Appl. Probab. 15 (2005), no. 2, 1433–1444

[8] Jacod, J.: *Calcul stochastique et problèmes de martingales.* Lecture Notes in Mathematics, 714. Springer, Berlin, 1979.

[9] Kunita, H. *Stochastic flows of diffeomorphisms and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, 24. Cambridge University Press, Cambridge, 1990.

[10] Kusuoka, S.; Stroock, D. : *Applications of the Malliavin calculus. II.* J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1985), no. 1, 1–76.

[11] Millet, A.; Nualart, D. and Sanz, M. *Integration by parts and time reversal for diffusion processes*, Ann. Probab. 17 (1989), no. 1, 208–238.

[12] Pardoux, E.:*Grossissement d'une filtration et retournement du temps d'une diffusion*, Séminaire de Probabilités, XX, 1984/85, 48–55, Lecture Notes in Math., 1204, Springer, Berlin, 1986.

[13] Revuz, D.; Yor, M.: *Continuous martingales and Brownian motion.* Third edition. Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin, 1999.